

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2006, 11:05:32 ; Search time 259 Seconds

(without alignments)
3418.680 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815

Sequence: 1 MELALCRWGLLALLPAGA.....TFKGTPTAEVPEYGLDVPV 1255

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : 1: uniprot_05.80.*

2: uniprot_sprot.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6815	100.0	1255	1 ERBB2_HUMAN	P04626 homo sapien
2	6304	92.5	1259	1 ERBB2_CANPA	O18733 canis famil
3	6007.5	88.2	1259	2 O8K3F5_RAT	P06494 rattus norv
4	6003	88.1	1257	1 ERBB2_RAT	P06494 rattus norv
5	5993.5	87.9	1254	1 ERBB2_MESAU	P06053 mesocricetu
6	5982.5	87.8	1305	2 O6ZPE0_MOUSE	O6ZPE0 mus musculu
7	4216	61.9	881	2 O8C0E7_MOUSE	O8C0E7 m mus muscu
8	3514.5	51.6	1275	2 O5EBY4_BRARE	O5EBY4 brachydanto
9	3256.5	47.8	711	2 O80Y89_MOUSE	O80Y89 mus musculu
10	3172	46.5	1209	2 O9GX70_RAT	O9GX70 rattus norv
11	3167	46.5	1210	1 EGFR_HUMAN	P00533 homo sapien
12	3152.5	46.3	1209	2 O8M1L8_PIG	O8M1L8 sus scrofa
13	3146	46.2	1210	1 EGFR_MOUSE	O01279 mus musculu
14	3146	46.2	1210	2 O5SVB8_MOUSE	O5SVB8 mus musculu
15	3143	46.1	1210	2 O6PPE8_MOUSE	O6PPE8 mus musculu
16	3045	44.7	1292	2 O4PLA5_CHICK	O4PLA5 gallus gall
17	3036	44.5	1276	2 O4PLA4_CHICK	O4PLA4 gallus gall
18	3006	44.1	1292	2 O6UZA6_RAT	O6UZA6 rattus norv
19	3004.5	44.1	1308	1 ERBB4_HUMAN	O51503 homo sapien
20	3004	44.1	1308	2 O6UZA9_RAT	O6UZA9 rattus norv
21	2989	43.9	1091	1 ERBB4_RAT	O504U8 homo sapien
22	2988.5	43.9	1091	2 O504U8_HUMAN	O504U8 homo sapien
23	2986.5	43.8	1191	2 O76ZP7_BRARE	O76ZP7 brachydanto
24	2977.5	43.7	1191	2 O6VOA3_BRARE	O6VOA3 brachydanto
25	2883.5	41.3	1209	2 O6XV8_XIPXI	O6XV8 xiphophorus
26	2857	40.9	1081	2 O59FL8_HUMAN	O59FL8 homo sapien
27	2758.5	40.5	1335	2 O4RG29_TESTG	O4RG29 tetrarodon n
28	2758	40.1	1167	2 O9YH40_XIPXI	O9YH40 xiphophorus
29	2735.5	40.1	1137	2 O9W6F6_CHICK	O9W6F6 gallus gall
30	2724.5	40.0	1167	1 XM8K_XIPXA	P13388 xiphophorus
31	2620.5	38.5	1094	2 O4RWY4_TESTNG	O4RWY4 tetrarodon n

32	2441.5	35.8	1342	1 ERBB3_HUMAN	P21860 homo sapien
33	2438.5	35.8	1342	2 O5RB22_PONPY	O5RB22 pongo pygma
34	2408	35.3	1339	2 O6B064_MOUSE	O6B064 mus musculu
35	2370.5	34.8	1339	1 ERBB3_RAT	O62799 rattus norv
36	2326	34.1	1328	2 P79754_FUGRU	P79754 fugu rubrip
37	2282	33.5	1429	2 O5G254_BRARE	O5G254 brachydanto
38	2209.5	32.4	1305	2 O8AW81_BRARE	O8AW81 brachydanto
39	2137.5	31.4	1344	2 O4RW26_TESTG	O4RW26 tetrarodon n
40	2102	30.8	892	2 O59EM4_HUMAN	O59EM4 homo sapien
41	2066	30.3	1547	2 O4RXV8_TESTG	O4RXV8 tetrarodon n
42	2062.5	30.3	1321	2 O7PENS_ANOGA	O7PENS anopheles g
43	2059.5	30.2	1375	2 O7FHU6_ANOGA	O7FHU6 anopheles g
44	2043.5	30.0	1433	2 O9BIH9_ANOGA	O9BIH9 anopheles g
45	2024.5	29.7	435	2 O6ZMM4_HUMAN	O6ZMM4 homo sapien

ALIGNMENTS

RESULT 1

ERBB2_HUMAN STANDARD: PRT: 1255 AA.

AC P04626; Q14256; Q6LDV1; Q9UMK4;

DT 13-AUG-1987 (Rel. 05, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Receptor tyrosine-protein kinase erbB-2 precursor (BC 2.7.1.112)

DE (p185erbB2) (C-erbB-2) (NEU proto-oncogene) (Tyrosine kinase-type cell

DE surface receptor HER2) (MIM 19).

GN Name=ERBB2; Synonyma=HER2, NEU, NGL;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;

OC Homo.

OC NCBI_Taxid=9606;

OC (1)

RP MEDLINE=86118663; PubMed=302577;

RA Yamamoto T., Ikawa S., Akizawa T., Semba K., Nomura N., Miyajima N.,

RA Saito T., Toyoshima K.;

RT "Similarity of protein encoded by the human c-erb-B-2 gene to

RT epidermal growth factor receptor.";

RL Nature 319:230-234(1986).

RL (2)

RX MEDLINE=86070181; PubMed=2999974;

RA Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J.,

RA Seeburg P.H., Libermann T.A., Schlessinger J., Francke U.,

RA Levinson A., Ullrich A.;

RT "Tyrosine kinase receptor with extensive homology to EGF receptor

RT shares chromosomal location with neu oncogene.";

RL Science 230:1132-1139(1985).

RL (3)

RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS CYS-452; VAL-655 AND

RP ALA-1170.

RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,

RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,

RA Schickelitz W.S., Sherwood J.K., Wittek L.A., Nickerson D.A.;

RT "NIH-SNPs, environmental genome project, NIBS B15478, Department

RT of Genome Sciences, Seattle, WA (URL: http://esg.gs.washington.edu).";

RL Submitted (DEC-2002) to the EMBL/Genbank/DBJ databases.

RL (4)

RP NUCLEOTIDE SEQUENCE OF 1-191.

RA MEDLINE=87286898; PubMed=3039351;

RA Tal M., King C.R., Kraus M.H., Ullrich A., Schlessinger J., Givol D.;

RT "Human HER2 (neu) promoter: evidence for multiple mechanisms for

RT transcriptional initiation.";

RL Mol. Cell. Biol. 7:2597-2601(1987).

RL (5)

RP NUCLEOTIDE SEQUENCE OF 737-1031.

RA MEDLINE=86016729; PubMed=2995967;

RA Semba K., Kamata N., Toyoshima K., Yamamoto T.;

RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-

RT erbB-1/epidermal growth factor-receptor gene and is amplified in a
RT human salivary gland adenocarcinoma".
CC Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).
CC [6]
RN NUCLEOTIDE SEQUENCE OF 832-909.
RP TISSUE=Mammary carcinoma;
RX MEDLINE=85372597; PubMed=2292089;
RA King C.R., Kraus M.H., Aaronson S.A.;
RT "Amplification of a novel v-erbB-related gene in a human mammary
RT carcinoma.";
RL Science 229:974-976(1985).
RN [7]
RN NUCLEOTIDE SEQUENCE OF 1081-1245, AND VARIANT ALA-1170.
RP MEDLINE=94000386; PubMed=8104414;
RX Sarker F.H., Ball D.B., Li Y.W., Criseman J.D.;
RA "Molecular cloning and sequencing of an intron of Her-2/neu (ERBB2)
RT gene.";
RL DNA Cell Biol. 12:611-615(1993).
RN [8]
RP IDENTIFICATION IN A COMPLEX WITH PIK3C2A AND EGFR, IDENTIFICATION IN A
RP COMPLEX WITH PIK3C2B AND EGFR, AND INTERACTION WITH PIK3C2B.
RX PubMed=10805725;
RA Arcaro A., Zvelebil M.J., Wallisch C., Ullrich A., Waterfield M.D.,
RA Domain J.;
RT "Class II phosphoinositide 3-kinases are downstream targets of
RT activated polypeptide growth factor receptors.";
PL Mol. Cell. Biol. 20:3817-3830(2000).
RN [9]
RP INTERACTION WITH PLXNB1.
RX PubMed=15210733; DOI=10.1083/jcb.200312094;
RA Swiercz J.M., Kuner R., Offermanns S.;
RT "Plaxin-B1/RhogEF-mediated RhoA activation involves the receptor
RT tyrosine kinase ErbB-2.";
RL J. Cell Biol. 165:869-880(2004).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 654-662 IN COMPLEX WITH H1A
RP AND BETA-2 MICROGLOBULIN.
RX MEDLINE=20062861; PubMed=10593938; DOI=10.1074/jbc.274.51.36422;
RA Kuhns J.J., Batalja M.A., Van S., Collins E.J.;
RT "Poor binding of a HER-2/neu epitope (GP) to H1A-A2.1 is due to a
RT lack of interactions with the center of the peptide.";
RL J. Biol. Chem. 274:36422-36427(1999).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.52 ANGSTROMS) OF 23-629 IN COMPLEX WITH FAB.
RX MEDLINE=22497871; PubMed=12610629; DOI=10.1038/nature01392;
RA Cho H.-S., Mason K., Ramyar K.X., Stanley A.M., Gabelli S.B.,
RA Denney D.W., Jr., Leahy D.J.;
RT "Structure of the extracellular region of HER2 alone and in complex
RT with the Herceptin Fab.";
RL Nature 421:756-760(2003).
RN [12]
RP VARIANTS VAL-654 AND VAL-655.
RX MEDLINE=93194196; PubMed=8095488;
RA Ehsani A., Low J., Wallace R.B., Wu A.M.;
RT "Characterization of a new allele of the human ERBB2 gene by allele-
RT specific competition hybridization.";
RL Genomics 15:426-429(1993).
CC -1- FUNCTION: Essential component of a neurogulin-receptor complex,
CC although neurogulins do not interact with it alone. GP30 is a
CC potential ligand for this receptor. Not activated by EGF, TGF-
CC alpha and amphiregulin.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein
CC tyrosine phosphate.
CC -1- SUBUNIT: Heterodimer with each of the other ERBB receptors
CC (potential). Interacts with PRKCAP (By similarity). Binds PLXNB1.
CC Part of a complex with EGFR and either PIK3C2A or PIK3C2B. May
CC interact with PIK3C2B when phosphorylated on Tyr-1196.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine residues
CC (By similarity).
CC -1- POLYMORPHISM: There are four alleles due to the variations in
CC positions 654 and 655. Allele B1 (Ile-654/Val-655) has a frequency
CC of 0.782; allele B2 (Ile-654/Val-655) has a frequency of 0.206;

CC allele B3 (Val-654/Val-655) has a frequency of 0.012.
CC -1- SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL, M11767, AAA35808.1, -, Genomic_DNA.
DR EMBL, M11761, AAA35808.1, JOINED, Genomic_DNA.
DR EMBL, M11762, AAA35808.1, JOINED, Genomic_DNA.
DR EMBL, M11763, AAA35808.1, JOINED, Genomic_DNA.
DR EMBL, M11764, AAA35808.1, JOINED, Genomic_DNA.
DR EMBL, M11765, AAA35808.1, JOINED, Genomic_DNA.
DR EMBL, M11766, AAA35808.1, JOINED, Genomic_DNA.
DR EMBL, M11730, AAA35808.1, -, mRNA.
DR EMBL, M12036, AAA35978.1, -, Genomic_DNA.
DR EMBL, AY208911, AAC18082.1, -, Genomic_DNA.
DR EMBL, X03363, CAA27060.1, -, mRNA.
DR EMBL, M16792, AAA58637.1, -, Genomic_DNA.
DR EMBL, M16789, AAA58637.1, JOINED, Genomic_DNA.
DR EMBL, M16790, AAA58637.1, JOINED, Genomic_DNA.
DR EMBL, M16791, AAA58637.1, JOINED, Genomic_DNA.
DR EMBL, L29395, AAA35809.1, -, Genomic_DNA.
DR EMBL, M95667, AAC37531.1, -, Unassigned_DNA.
DR PIR, A24571, A24571.
DR PDB, 1N8Z, X-ray; C-23-629.
DR PDB, 1OVG, Model; A=737-1031.
DR PDB, 1S78, X-ray; A/B=23-646.
DR Ensembl, ENSG00000141736, Homo sapiens.
DR HGNC, HGNC:3430, ERBB2.
DR MIM, 164870, -;
DR GO, GO:0016021, C:Integral to membrane; NAS.
DR GO, GO:0005866, C:Plasma membrane; NAS.
DR GO, GO:0005006, F:epidermal growth factor receptor activity; NAS.
DR GO, GO:0043125, F:ErbB-3 class receptor binding; TAS.
DR GO, GO:0046982, F:Protein heterodimerization activity; NAS.
DR GO, GO:0004716, F:Receptor signaling protein tyrosine kinase . . . ; TAS.
DR GO, GO:0008283, P:cell proliferation; TAS.
DR GO, GO:0007507, P:heart development; TAS.
DR GO, GO:0030879, P:mammary gland development; TAS.
DR GO, GO:0007339, P:neurogenesis; TAS.
DR GO, GO:0048015, P:phosphoinositide-mediated signaling; NAS.
DR GO, GO:0006468, P:protein amino acid phosphorylation; TAS.
DR GO, GO:0045765, P:regulation of angiogenesis; NAS.
DR GO, GO:007169, P:transmembrane receptor protein tyrosine kin. . . ; NAS.
DR InterPro, IPR000494, EGFR_L.
DR InterPro, IPR006211, Furin-like.
DR InterPro, IPR006212, Furin repeat.
DR InterPro, IPR000719, Prot. Kinase.
DR InterPro, IPR001245, Tyr_Kinase.
DR InterPro, IPR008266, Tyr_Kinase_AS.
DR InterPro, IPR004019, Furin-like, 1.
DR Pfam, PF00757, Furin-like, 1.
DR Pfam, PF01030, Recep_L_domain, 2.
DR Pfam, PF02757, YLP, 2.
DR PRINTS, PR00109, TYRKINASE.
DR PRODOM, PD000001, Prot_Kinase, 1.
DR SMART, SM00261, FU, 3.
DR SMART, SM00219, TYRK, 1.
DR PROSITE, PS00107, PROTEIN_KINASE_ATP, 1.
DR PROSITE, PS50011, PROTEIN_KINASE_DOM, 1.
DR PROSITE, PS00109, PROTEIN_KINASE_TYR, 1.
DR 3D-structure, ATP-binding, Glycoprotein; Kinase; Multigene family;
DR Nucleotide-binding, Phosphorylation; Polymorphism; Receptor; Signal;
DR Transferrase; Transmembrane; Tyrosine-protein kinase.
DR SIGNAL, 1, 21, Potential.
DR CHAIN, 22, 1255, Receptor tyrosine-protein kinase erbB-2.
DR TOPO_DOM, 22, 652, Extracellular (Potential).
DR TRANSMEM, 653, 675, Potential.
DR TOPO_DOM, 676, 1255, Cytoplasmic (Potential).

FT	DOMAIN	720	987	Protein kinase.
FT	NP_BIND	726	734	ATP (By similarity).
Query Match				
Beat Local Similarity 100.0%; Score 6815; DB 1; Length 1255;				
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	MELALACRMGLLALLP	PGASNOVCTGDMKRLRPASPEYTHDMRLHYOGGQVVGSL	60
DB	1	MELALACRMGLLALLP	PGASNOVCTGDMKRLRPASPEYTHDMRLHYOGGQVVGSL	60
QY	61	ELTYLPYTNASLFL	ODIQEVGVYLIANQVQVPLQRLIRVGTQLFEDNYALAVLDNG	120
DB	61	ELTYLPYTNASLFL	ODIQEVGVYLIANQVQVPLQRLIRVGTQLFEDNYALAVLDNG	120
QY	121	DELANTTPTVTGAS	PGGLRELQLRLTEILKGGVLIQRNPQLCYODTILMKDIFHKNNQLA	180
DB	121	DELANTTPTVTGAS	PGGLRELQLRLTEILKGGVLIQRNPQLCYODTILMKDIFHKNNQLA	180
QY	181	LTLIDPNRBRACH	PGSPKMGSRMGSESDCOSLRTVCAGGACAKGRLPTDCHEQC	240
DB	181	LTLIDPNRBRACH	PGSPKMGSRMGSESDCOSLRTVCAGGACAKGRLPTDCHEQC	240
QY	241	AAAGCTGPKHSDCL	ACIAHFHNSGICELHCPALVTYNTDYESMNPREGRYTFGASCVTACP	300
DB	241	AAAGCTGPKHSDCL	ACIAHFHNSGICELHCPALVTYNTDYESMNPREGRYTFGASCVTACP	300
QY	301	YNYLSTDVGSCT	LVCPHLNQEVTAEDGTORCEKSPCARVCYGLGMEHLREYRATTSAN	360
DB	301	YNYLSTDVGSCT	LVCPHLNQEVTAEDGTORCEKSPCARVCYGLGMEHLREYRATTSAN	360
QY	361	IOEPFGCKKIFGSL	AFLPSPGDPASNTAPLOPBLQVFEYTLBEITGLYISAMPDSAP	420
DB	361	IOEPFGCKKIFGSL	AFLPSPGDPASNTAPLOPBLQVFEYTLBEITGLYISAMPDSAP	420
QY	421	DLVSFQNLQVIR	GIILHNGAVSLTLOGLGISWLGRLSLRELSGSLIHHNTHLCFVHTV	480
DB	421	DLVSFQNLQVIR	GIILHNGAVSLTLOGLGISWLGRLSLRELSGSLIHHNTHLCFVHTV	480
QY	481	PMDQLFRNPHQALL	HTANRPEBECVGEGLACHQLCARHGCMGPPQVCNCSQFLRGQRC	540
DB	481	PMDQLFRNPHQALL	HTANRPEBECVGEGLACHQLCARHGCMGPPQVCNCSQFLRGQRC	540
QY	541	VEECVTLQGLPRE	VYNAHCLPCHPECPONSGVTCFGEBAQCVACAHYKDPFVAVAC	600
DB	541	VEECVTLQGLPRE	VYNAHCLPCHPECPONSGVTCFGEBAQCVACAHYKDPFVAVAC	600
QY	601	PGSVPRDLSPYMI	WKPRDEGACQPCPINCTHSQVDLDDKGCPEAQASPLTSISAVVG	660
DB	601	PGSVPRDLSPYMI	WKPRDEGACQPCPINCTHSQVDLDDKGCPEAQASPLTSISAVVG	660
QY	661	ILLVVLGVVFGILL	KRQOKIRKTYMRLLOETELVEBPLTPSGAMPNOQMRILKETEL	720
DB	661	ILLVVLGVVFGILL	KRQOKIRKTYMRLLOETELVEBPLTPSGAMPNOQMRILKETEL	720
QY	721	RKYKVLGSGAFGV	YVYGIWIPDGENYKIVAKVLEBNTSPRANKELDEAYVMAGVSP	780
DB	721	RKYKVLGSGAFGV	YVYGIWIPDGENYKIVAKVLEBNTSPRANKELDEAYVMAGVSP	780
QY	781	YYSRLLGICLTST	VQLTQMLMPYGLLDHYRENRGLSGODLLNMCQIAKGSYLEDEV	840
DB	781	YYSRLLGICLTST	VQLTQMLMPYGLLDHYRENRGLSGODLLNMCQIAKGSYLEDEV	840
QY	841	LVHRDLAANVIL	VKSFNHYKIDTFGLARLLIDETETSHADGKVPITKMALESILERRRT	900
DB	841	LVHRDLAANVIL	VKSFNHYKIDTFGLARLLIDETETSHADGKVPITKMALESILERRRT	900
QY	901	HOSDWSYGVTV	WEIMTFGAKPYDGI PARBIPLLEKGRLLPPICTIDYVMIMKCM	960
DB	901	HOSDWSYGVTV	WEIMTFGAKPYDGI PARBIPLLEKGRLLPPICTIDYVMIMKCM	960
QY	961	IDSECRPREREL	VSFSESRMARDPQRFVYIQNEIDLGPASPLDSTFYRSLLDEDDMGDLVDA	1020
DB	961	IDSECRPREREL	VSFSESRMARDPQRFVYIQNEIDLGPASPLDSTFYRSLLDEDDMGDLVDA	1020

DB	961	IDSECRPREREL	VSFSESRMARDPQRFVYIQNEIDLGPASPLDSTFYRSLLDEDDMGDLVDA	1020
QY	1021	EEYLVPOGFCPC	PDAPAGAGVHHRHSSTRSGGDLTLGLRPSSEEPAPSPPLAPSBG	1080
DB	1021	EEYLVPOGFCPC	PDAPAGAGVHHRHSSTRSGGDLTLGLRPSSEEPAPSPPLAPSBG	1080
QY	1081	AGSDVFDGGLG	MAKGLQSLPTHDPSPLQRYSEDPVPLPSEETDGYVAPLTCSPQPEYV	1140
DB	1081	AGSDVFDGGLG	MAKGLQSLPTHDPSPLQRYSEDPVPLPSEETDGYVAPLTCSPQPEYV	1140
QY	1141	NOPDVRPOPSP	REBGLPAPRAGATLERPKTLSSGKNGVMDVAFGAVENPEYLTQ	1200
DB	1141	NOPDVRPOPSP	REBGLPAPRAGATLERPKTLSSGKNGVMDVAFGAVENPEYLTQ	1200
QY	1201	GGAAPOPHP	PPAFSPAFDLYYMDOPPERGAPSTFKOTPTAENPEYLGDLVVP	1255
DB	1201	GGAAPOPHP	PPAFSPAFDLYYMDOPPERGAPSTFKOTPTAENPEYLGDLVVP	1255

RESULT 2

ERBB2	CANPA	STANDARD;	PRT; 1259 AA.
ID	ERBB2	CANFA	
AC	018735;		
DT	25-OCT-2004 (rel. 45, Created)		
DT	25-OCT-2004 (rel. 45, Last sequence update)		
DT	13-SEP-2005 (rel. 48, Last annotation update)		
DE	Receptor tyrosine-protein kinase erbB-2 precursor (EC 2.7.1.112)		
DE	(p15terb2) (C-erbB-2).		
GN	Name=ERBB2;		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;		
OC	Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;		
OC	Canis.		
OX	NCBI_TaxID=9615;		
RN	(1)		
RP	NUCLEOTIDE SEQUENCE.		
RC	TISSUE=Mammary gland;		
RA	Yokota H.;		
RT	"cDNA cloning of erbB-2 from canine mammary gland."		
RL	Submitted (OCT-1997) to the EMBL/Genbank/DBD databases.		
CC	- FUNCTION: Essential component of a neurotrophin-receptor complex,		
CC	although neurotrophins do not interact with it alone. GP30 is a		
CC	potential ligand for this receptor. Not activated by EGF, TGF-		
CC	alpha and amphiregulin (By similarity).		
CC	- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein		
CC	tyrosine phosphate.		
CC	- SUBUNIT: Heterodimer with each of the other ERBB receptors		
CC	(potential). Interacts with PRKCAP. Binds PLXNB1. Part of a		
CC	complex with EGFR and either PIK3C2A or PIK3C2B. May interact with		
CC	PIK3C2B when phosphorylated on Tyr-1200 (By similarity).		
CC	- SUBCELLULAR LOCATION: Type I membrane protein.		
CC	- PTM: Ligand-binding increases phosphorylation on tyrosine residues		
CC	(By similarity).		
CC	- SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor		
CC	subfamily.		

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; AB008451; BAA23127.1; -; mRNA.

HSSP; P04626; 1N8Z.

DR SMR; O18735; 23-628.

DR Ensembl; ENSGAFG0000016351; Canis familiaris.

DR InterPro; IPR000494; EGFR_L.

DR InterPro; IPR006211; Furin-like.

DR InterPro; IPR006212; Furin repeat.

DR InterPro; IPR000719; Prot_Kinase.

DR InterPro; IPR001245; Tyr_kinase.

DR InterPro; IPR008266; Tyr_kinase_AS.

DR InterPro; IPR004019; YLP_motif.

DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR Pfam: PF0257; VLP; 2.
DR PRINTS: PR00109; TYRKINSE.
DR ProDom: PD000001; Proc_Kinase; 1.
DR SMART: SM00261; FU; 3.
DR SMART: SM00219; TyKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Glycoprotein; Kinase; Multigene family;
KW Nucleotide-binding; Phosphorylation; Receptor; Signal; Transferase;
KW Transmembrane; Tyrosine-protein kinase.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 1259 Receptor tyrosine-protein kinase erbB-2.
FT TOPO_DOM 23 653 Extracellular (Potential).
FT TRANSMEM 654 674 Potential.
FT TOPO_DOM 675 1259 Cytoplasmic (Potential).
FT DOMAIN 719 986 Protein kinase.
FT NP_BIND 725 733 ATP (By similarity).
FT REGION 1199 1201 Interaction with PIK3C2B (By similarity).
FT COMBINS 192 268 Cys-rich.
FT COMBINS 1101 1219 Pro-rich.
FT ACT_SITE 844 844 By similarity.
FT BINDING 752 752 ATP (By similarity).
FT MOD_RES 1138 1138 Phosphotyrosine (by autocatalysis) (By similarity).
FT MOD_RES 1200 1200 Phosphotyrosine (Potential).
FT MOD_RES 1252 1252 Phosphotyrosine (by autocatalysis) (By similarity).
FT CARBOHYD 68 68 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 259 259 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 421 421 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 529 529 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 570 570 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 628 628 N-linked (GlcNAc...) (Potential).
FT DISULFID 195 204 By similarity.
FT DISULFID 199 212 By similarity.
FT DISULFID 220 227 By similarity.
FT DISULFID 224 235 By similarity.
FT DISULFID 236 244 By similarity.
FT DISULFID 240 252 By similarity.
FT DISULFID 255 264 By similarity.
FT DISULFID 268 295 By similarity.
FT DISULFID 299 311 By similarity.
FT DISULFID 315 331 By similarity.
FT DISULFID 334 338 By similarity.
FT DISULFID 511 519 By similarity.
FT DISULFID 514 527 By similarity.
FT DISULFID 530 539 By similarity.
FT DISULFID 543 559 By similarity.
FT DISULFID 562 575 By similarity.
FT DISULFID 586 583 By similarity.
FT DISULFID 595 595 By similarity.
FT DISULFID 599 622 By similarity.
FT DISULFID 625 633 By similarity.
FT DISULFID 629 641 By similarity.
SQ SEQUENCE 1259 AA; 137991 MW; E37364D9CAACD46 CRC64;

Query Match 92.5%; Score 6304; DB 1; Length 1259;
Best Local Similarity 92.2%; Pred. No. 1.6e-312;
Matches 1162; Conservative 38; Mismatches 54; Indels 6; Gaps 2;

QY 1 MELALCRMGILLALLPFGAASVOYCTGDMKLRIPASPTHLLDMLRHLVYGGCCVVGQNL 60
DB 1 MELAWCMWGLLALLPSGAAGTQCTGDMKLRIPASPTHLLDMLRHLVYGGCCVVGQNL 60
QY 61 ETTYLPFTNASLFLDIOEVQGVLIANQVQVFLQRLIYVGTQLFEDNYALAVLDNG 120
DB 61 ETTYLPFTNASLFLDIOEVQGVLIANQVQVFLQRLIYVGTQLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTTIIMKDTIFKNNOLA 180
DB 121 DPLEGGIPAPGAAGGLRELQRLSLTEILKGGVLIQRNPOLCHODTTIMKDVIFKNNOLA 180

DB 121 DPLEGGIPAPGAAGGLRELQRLSLTEILKGGVLIQRNPOLCHODTTIMKDVIFKNNOLA 180
QY 181 LTLIDTNRSPACHPSCPMCKGSRGCGSSBDCQSLTRTVACGACGKGPLPTDCCHQC 240
DB 181 LTLIDTNRSPACHPSCPMCKGSRGCGSSBDCQSLTRTVACGACGKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLALHPNHSIGICELHPALVTYNTDTFESMPNDEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLALHPNHSIGICELHPALVTYNTDTFESMPNDEGRYTFGASCVTACP 300
QY 301 YNYLSTVGSCTTVCPLINQEVTAEDGTQRECKSKPCARVCYGLGMEHLREVAIVSAN 360
DB 301 YNYLSTVGSCTTVCPLINQEVTAEDGTQRECKSKPCARVCYGLGMEHLREVAIVSAN 360
QY 361 IQEPAGCKKIFGSLAFLPESFPGDPAANTAPLOEQLQVFETLEITGYLYISAMPDLP 420
DB 361 IQEPAGCKKIFGSLAFLPESFPGDPAANTAPLOEQLQVFETLEITGYLYISAMPDLP 420
QY 421 DLSVFQNLQVIRGRIILNNGAVSLTLQGLGISWGLRSLREISGLALIHNTHLCPVHTV 480
DB 421 NLSVFQNLQVIRGRIILNNGAVSLTLQGLGISWGLRSLREISGLALIHNTHLCPVHTV 480
QY 481 PMDQLFNNPHQALHTNRPDECVGEGGLACHQLCARGHCWGPRPTCVNCSQFLRGQC 540
DB 481 PMDQLFNNPHQALHTNRPDECVGEGGLACHQLCARGHCWGPRPTCVNCSQFLRGQC 540
QY 541 VEEGRVQLGPREYVYNAHCLPCHPECOPQNGSYTCGPEADQCVACAHYDPPFCVARC 600
DB 541 VEEGRVQLGPREYVYNAHCLPCHPECOPQNGSYTCGPEADQCVACAHYDPPFCVARC 600
QY 540 VEEGRVQLGPREYVYNAHCLPCHPECOPQNGSYTCGPEADQCVACAHYDPPFCVARC 599
DB 540 VEEGRVQLGPREYVYNAHCLPCHPECOPQNGSYTCGPEADQCVACAHYDPPFCVARC 599
QY 601 PSQVYKPDLSYMPYKPFDEBGACQPCPINCCHSCVDLDDKCCPABORASPLTSIIAVVG 660
DB 601 PSQVYKPDLSYMPYKPFDEBGACQPCPINCCHSCVDLDDKCCPABORASPLTSIIAVVG 660
QY 600 PSQVYKPDLSYMPYKPFDEBGACQPCPINCCHSCVDLDDKCCPABORASPLTSIIAVVG 659
DB 600 PSQVYKPDLSYMPYKPFDEBGACQPCPINCCHSCVDLDDKCCPABORASPLTSIIAVVG 659
QY 661 ILVAVVGLVGLIKRQOKIRKYTMRLLOEELVBPRLPSGAMPQAOQMLKETEL 720
DB 661 ILVAVVGLVGLIKRQOKIRKYTMRLLOEELVBPRLPSGAMPQAOQMLKETEL 720
QY 721 RKVAVLSSGAFYVYKGIWIPDGENVKIPVAIKYLRNTSPKANKELIDEAIVVAGVSP 780
DB 721 RKVAVLSSGAFYVYKGIWIPDGENVKIPVAIKYLRNTSPKANKELIDEAIVVAGVSP 780
QY 720 RKVAVLSSGAFYVYKGIWIPDGENVKIPVAIKYLRNTSPKANKELIDEAIVVAGVSP 779
DB 720 RKVAVLSSGAFYVYKGIWIPDGENVKIPVAIKYLRNTSPKANKELIDEAIVVAGVSP 779
QY 781 YVSRLLGICLTSTVQVLTQMPYGCCLLDHVENRGLSGODLAWCMQIAGKSYLBYR 840
DB 781 YVSRLLGICLTSTVQVLTQMPYGCCLLDHVENRGLSGODLAWCMQIAGKSYLBYR 840
QY 780 YVSRLLGICLTSTVQVLTQMPYGCCLLDHVENRGLSGODLAWCMQIAGKSYLBYR 839
DB 780 YVSRLLGICLTSTVQVLTQMPYGCCLLDHVENRGLSGODLAWCMQIAGKSYLBYR 839
QY 841 LVHBDLAARNLVKSPNHVKITDFGLARLLDIDETVHADGKVPKIMMALESLIRRF 900
DB 841 LVHBDLAARNLVKSPNHVKITDFGLARLLDIDETVHADGKVPKIMMALESLIRRF 900
QY 901 HQSDVMSGYVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPORPICIIVYMMIVKCM 960
DB 901 HQSDVMSGYVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPORPICIIVYMMIVKCM 960
QY 900 HQSDVMSGYVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPORPICIIVYMMIVKCM 959
DB 900 HQSDVMSGYVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPORPICIIVYMMIVKCM 959
QY 961 IDSECRPFRELVEFSRMAADPQRFVYIIONEDLGPASPLDSTYRSLLEDMDGLVDA 1020
DB 961 IDSECRPFRELVEFSRMAADPQRFVYIIONEDLGPASPLDSTYRSLLEDMDGLVDA 1020
QY 960 IDSECRPFRELVEFSRMAADPQRFVYIIONEDLGPASPLDSTYRSLLEDMDGLVDA 1019
DB 960 IDSECRPFRELVEFSRMAADPQRFVYIIONEDLGPASPLDSTYRSLLEDMDGLVDA 1019
QY 1021 EBYLVPOQGFPCRPAPGAGGMVHNHRSSTRSGGDLTDLGPSEEARSPPLAPSEG 1080
DB 1021 EBYLVPOQGFPCRPAPGAGGMVHNHRSSTRSGGDLTDLGPSEEARSPPLAPSEG 1080
QY 1020 EBYLVPOQGFPCRPAPGAGGMVHNHRSSTRSGGDLTDLGPSEEARSPPLAPSEG 1079
DB 1020 EBYLVPOQGFPCRPAPGAGGMVHNHRSSTRSGGDLTDLGPSEEARSPPLAPSEG 1079
QY 1081 AGSDVFDGDLGMGAAGKLGSLPTDPSPLQRYSDPTPLPSEYDGVAPLPTCSPOPEYV 1140
DB 1081 AGSDVFDGDLGMGAAGKLGSLPTDPSPLQRYSDPTPLPSEYDGVAPLPTCSPOPEYV 1140
QY 1080 AGSDVFDGDLGMGAAGKLGSLPTDPSPLQRYSDPTPLPSEYDGVAPLPTCSPOPEYV 1139
DB 1080 AGSDVFDGDLGMGAAGKLGSLPTDPSPLQRYSDPTPLPSEYDGVAPLPTCSPOPEYV 1139
QY 1141 NQPDVRRPQSPREGRPLPAPAPACATLBR-----PKTISPKNGVYDVPFAGAVENPE 1195
DB 1141 NQPDVRRPQSPREGRPLPAPAPACATLBR-----PKTISPKNGVYDVPFAGAVENPE 1195
QY 1140 NQPEWVQPPPLALGSPRPSPRPAATLERPKTLSPKTLSPKNGVYDVPFAGAVENPE 1199
DB 1140 NQPEWVQPPPLALGSPRPSPRPAATLERPKTLSPKTLSPKNGVYDVPFAGAVENPE 1199
QY 1196 YLTPOGGAAPQHPHPAPAFSAFDNLVYWDQDPPRPGAPSPSPFKGPTAPNEPEYGLDVPV 1255
DB 1196 YLTPOGGAAPQHPHPAPAFSAFDNLVYWDQDPPRPGAPSPSPFKGPTAPNEPEYGLDVPV 1255
QY 1200 YLAPRGRAPQHPHPAPAFSAFDNLVYWDQDPPRPGAPSPSPFKGPTAPNEPEYGLDVPV 1259
DB 1200 YLAPRGRAPQHPHPAPAFSAFDNLVYWDQDPPRPGAPSPSPFKGPTAPNEPEYGLDVPV 1259


```
RESULT 3
Q8K3F9 RAT PRELIMINARY; PRT; 1259 AA.
ID Q8K3F9 RAT PRELIMINARY; PRT; 1259 AA.
AC Q8K3F9;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Neu proteinocytocytin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BDIX;
RA Watson P. A., Kim K., Chen K.-S., Gould M.N.;
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY116182; AAM50093.1; -; mRNA.
DR HSSP; P06494; 1N8Y.
DR SMR; Q8K3F9; 26-633.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR004019; YLP motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 4.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 1259 AA; 139101 MW; B724BDSCC3AB953 CRC64;

Query Match 88.2%; Score 6007.5; DB 2; Length 1259;
Best Local Similarity 88.1%; Pred. No. 2e-297;
Matches 1106; Conservative 49; Mismatches 100; Indels 1; Gaps 1;

QY 1 MELIALCGMGLTLLALPGCASTGYCTGDMKRLPASPEETHLMDLRHLHYOCQVVGNTL 60
DB 4 MELAWCMGFLTALPPEIAGTQCTGDMKRLPASPEETHLMDLRHLHYOCQVVGNTL 63
QY 61 ELTYLPTNASLSFLDIOEVQGYVLIANQVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
DB 64 ELTYVPAVASLSFLDIOEVQGYVLIANQVQVPLQRLIRVGTQLFEDNYALAVLDNR 123
QY 121 DPLANTTPTVT-GASFGGLRELOLRSLITELKGVLIQRNPOLCYODTIIMKQIFKKNQL 179
DB 124 DPODVVAASSTPGRTPEGRLRELQRLSLITELKGVLIQRNPOLCYODTMVLMKQVFRKNQL 183
QY 180 ALTLDTNRSRACHPCSPWCKGRGSGESSSDCOSLTRVAGGACGARKGPLPTDCCHQ 239
DB 184 APVDIDTNRSRACHPCSPWCKGRGSGESSSDCOSLTRVAGGACGARKGPLPTDCCHQ 243
QY 240 CAAGCTGPGHSDCLACLAHNSGICELACPALVYNTDTFESMPNPEGRTFGASCVTAC 299
DB 244 CAAGCTGPGHSDCLACLAHNSGICELACPALVYNTDTFESMPNPEGRTFGASCVTTC 303
```

```
QY 300 PYNLTSDVGSCTVLCPLAHNOEVTAEDETQCEKSKPCARVCYGLGMEHLREVAVTSA 359
DB 304 PYNLTSDVGSCTVLCPLAHNOEVTAEDETQCEKSKPCARVCYGLGMEHLREVAVTSD 363
QY 360 NIOEPAGCKKI FGSIAFLPESFDGDPASNTVPLQPEQLQVETTELITGYLISAMPSSL 419
DB 364 NIOEPAGCKKI FGSIAFLPESFDGDPASNTVPLQPEQLQVETTELITGYLISAMPSSL 423
QY 420 PDLVSFQNLQVIRGILHNGAVSLTLQGLGSIWGLRSLREISGSLALIHNTHLCFVHT 479
DB 424 PDLVSFQNLQVIRGILHNGAVSLTLQGLGSIWGLRSLREISGSLALIHNTHLCFVHT 483
QY 480 VPMQDLFNNPPIQALHTANRPEDECEVGLACHQICARHGMCPGPTQCVNCSQELRQOE 539
DB 484 VPMQDLFNNPPIQALHTANRPEDECEVGLACHQICARHGMCPGPTQCVNCSQELRQOE 543
QY 540 CVEECRVQLGLPREYVNAHRLCPHPECQOPONGSYTCGPEADQVACAHYKDPFCYAR 599
DB 544 CVEECRVWKGLPREYVSDKRLCPHPECQOPONSBTCPGSEADQVACAHYKSSCVAR 603
QY 600 CPSGVKPDLSYMPIMKPPDEEGACQPCPINCTHSCVDLDDRCPEAQRASPLTSLISAV 659
DB 604 CPSGVKPDLSYMPIMKPPDEEGACQPCPINCTHSCVDLDDRCPEAQRASPLTSLISAV 663
QY 660 GILVAVLVGVFGLIKRQOKIRKYTRRLQETELVEPLTPSGAMPNOQMRLKETE 719
DB 664 GVLFLFLIVVVGILIKRRQOKIRKYTRRLQETELVEPLTPSGAMPNOQMRLKETE 723
QY 720 LRKVVLTSSGAPGVYKGIWIPDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAGVGS 779
DB 724 LRKVVLTSSGAPGVYKGIWIPDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAGVGS 783
QY 780 PYVSRLLGICLTSTVQVLTOLMPYGCILDHRENRGRIGSODLQWQIAKMSYLEDV 839
DB 784 PYVSRLLGICLTSTVQVLTOLMPYGCILDHRENRGRIGSODLQWQIAKMSYLEDV 843
QY 840 RLVRHDLAARVNLVKS PNHVKITDFGLARLIDIDETEHADGKVPYIKMMALESILRRRF 899
DB 844 RLVRHDLAARVNLVKS PNHVKITDFGLARLIDIDETEHADGKVPYIKMMALESILRRRF 903
QY 900 THQSDVMSYGVTWELMTFGAKPYDGI PAREIPDLLEKGERLPOPPCTIDVYMTMKW 959
DB 904 THQSDVMSYGVTWELMTFGAKPYDGI PAREIPDLLEKGERLPOPPCTIDVYMTMKW 963
QY 960 MIDSECRFPRELVESEFRMARDDPORFVYIQNEDIGPASPILDSTYRSLSLEDDMDGLVD 1019
DB 964 MIDSECRFPRELVESEFRMARDDPORFVYIQNEDIGPASPILDSTYRSLSLEDDMDGLVD 1023
QY 1020 ABEYLVPOQGFPCPPAGAGGMVHRHRSSTRSGGDLTLGLEPSEEEAPRSLAPSE 1079
DB 1024 ABEYLVPOQGFPCPPAGAGGMVHRHRSSTRSGGDLTLGLEPSEEEAPRSLAPSE 1083
QY 1080 GAGSDVPFDGDLQMAKGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEY 1139
DB 1084 GAGSDVPFDGDLQMAKGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEY 1143
QY 1140 VNQDPVRQPPSPRRGRLPAARPAATLIERPTLISPGKNGVYKQVPAFGAENENEXYLT 1199
DB 1144 VNQDPVRQPPSPRRGRLPAARPAATLIERPTLISPGKNGVYKQVPAFGAENENEXYLT 1203
QY 1200 OGGAAPQHPAPPAPAFDNLVYMDQPPBKGAPSTFKGTPTAENPEYLGIDVY 1255
DB 1204 REGTASPPHPPAPAFDNLVYMDQNSSGQPPSNFEGTPTAENPEYLGIDVY 1259

RESULT 4
ERBB2 RAT STANDARD; PRT; 1257 AA.
ID ERBB2 RAT STANDARD; PRT; 1257 AA.
AC P06494; Q6P732;
DT 01-JAN-1988 (Rel. 06, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Receptor tyrosine-protein kinase erbB-2 precursor (EC 2.7.1.112)
```

(p185erbB2) (C-erbB-2) (NEU proto-oncogene) (Epidermal growth factor receptor-related protein).
 GN Name=Erbb2; Synonyms=Neu;
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; OC Muridae; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Neuroblastoma;
 RX MEDLINE=86116662; PubMed=3945311;
 RA Bargmann C.I., Hung M.-C., Weinberg R.A.;
 RT "The neu oncogene encodes an epidermal growth factor receptor-related protein."
 RL Nature 319:226-230(1986).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC TISSUE=Prostate;
 GN NIH - Mammalian Gene Collection (MGC) project;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 RP [3]
 RX NUCLEOTIDE SEQUENCE OF 634-699.
 RA Masui T., Mann A.M., Macatee T.L., Garland E.M., Okamura T., Smith R.A., Cohen S.W.;
 RT "Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no mutation in urinary bladder carcinomas induced by N-butyl-4-thiazolylformamide or N-methyl-N-nitrosourea."
 RL Carcinogenesis 12:1975-1978(1991).
 RN [4]
 RP NUCLEOTIDE SEQUENCE OF 852-905.
 RC TISSUE=Sciatic nerve;
 RX MEDLINE=9122560; PubMed=2025425; DOI=10.1016/0896-6273(91)90167-X; Lai C., Lemke G.;
 RT "An extended family of protein-tyrosine kinase genes differentially expressed in the vertebrate nervous system."
 RL Neuron 6:691-704(1991).
 RN [5]
 RP STRUCTURE BY NMR OF 650-668.
 RX MEDLINE=92155181; PubMed=1346763; Gullick W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D., Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
 RT "Three dimensional structure of the transmembrane region of the proto-oncogenic and oncogenic forms of the neu protein."
 RL EMBO J. 11:43-48(1992).
 CC -1- FUNCTION: Essential component of a neurogulin-receptor complex, although neurogulin do not interact with it alone. GP30 is a potential ligand for this receptor. Not activated by EGF, TGF-alpha and amphiregulin.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein tyrosine phosphate.
 CC -1- SUBUNIT: Heterodimer with each of the other ERBB receptors. The constructively activated oncogenic variant forms a homodimer. Interacts with PRKCAP and PLANK. Part of a complex with EGFR and either PIK3CA or PIK3CB. May interact with PIK3CB when phosphorylated on Tyr-1198 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: Ligand-binding increases phosphorylation on tyrosine residues (By similarity).
 CC -1- SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor subfamily.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC -----
 CC EMBL; X03362; CA27059.1; ALT_INIT; mRNA.
 CC DR EMBL; BC061863; AAH61863.1; ALT_INIT; mRNA.
 CC PIR; A24562; TVRTNU.

DR PDB; 1N8Y; X-ray; C=23-631.
 DR Ensembl; ENSRNOG0000006450; Rattus norvegicus.
 DR RCG; 2561; Erbb2.
 DR GO; GO:0045595; P:regulation of cell differentiation; TAS.
 DR GO; GO:0042127; P:regulation of cell proliferation; TAS.
 DR InterPro; IPR000464; EGRF_L.
 DR InterPro; IPR006211; Furin-like.
 DR InterPro; IPR006212; Furin_repeat.
 DR InterPro; IPR000719; Prot_Kinase.
 DR InterPro; IPR001245; Tyr_Kinase.
 DR InterPro; IPR008266; Tyr_Kinase_AS.
 DR InterPro; IPR004019; YLP_motif.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF01030; Recep_L_domain; 2.
 DR Pfam; PF02757; YLP; 2.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00261; FU; 4.
 DR SMART; SM00219; TYRC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 KW 3D-structure; ATP-binding; Disease mutation; Glycoprotein; Kinase; Multigene family; Nucleotide-binding; Phosphorylation; Proto-oncogene; Receptor; Signal; Transferase; Transmembrane; Tyrosine-protein kinase.
 KW SIGNAL.
 FT CHAIN 1..21
 FT SIGNAL 22..1257
 FT TOPO_DOM 22..654
 FT TOSMGM 655..677
 FT TOPO_DOM 678..1257
 FT DOMAIN 722..989
 FT NP_BIND 728..736
 FT REGION 1197..1199
 FT COMPBIAS 159..369
 FT ACT_SITE 473..646
 FT ACT_SITE 847..847
 FT BINDING 755..755
 FT MOD_RES 1141..1141
 FT MOD_RES 1198..1198
 FT MOD_RES 1250..1250
 FT CARBOHYD 68..68
 FT CARBOHYD 188..188
 FT CARBOHYD 260..260
 FT CARBOHYD 532..532
 FT CARBOHYD 573..573
 FT CARBOHYD 631..631
 FT DISULFID 196..205
 FT DISULFID 200..213
 FT DISULFID 221..228
 FT DISULFID 225..236
 FT DISULFID 237..245
 FT DISULFID 241..253
 FT DISULFID 256..265
 FT DISULFID 269..296
 FT DISULFID 300..312
 FT DISULFID 316..332
 FT DISULFID 335..339
 FT DISULFID 513..522
 FT DISULFID 517..530
 FT DISULFID 533..542
 FT DISULFID 546..562
 FT DISULFID 565..578
 FT DISULFID 569..586
 FT DISULFID 589..598
 FT DISULFID 602..625
 FT DISULFID 628..636
 FT DISULFID 632..644
 FT DISULFID 661..661
 FT VARIANT 632..661
 FT CONFLICT 145..145
 FT CONFLICT 505..509
 FT SEQUENCE 1257 AA; 138832 MW; 6129264583011402 CRC64;

Query Match 88.1%; Score 6003; DB 1; Length 1257;
Best Local Similarity 88.0%; Pred. No. 3,5e-297;
Matches 1106; Conservative 49; Mismatches 100; Indels 2; Gaps 2;

```

QY 1 MELAALCRWGLLLALLPPGAASSTQVCTGDMKRLRLPASPEHLDMLRLHLYGCGQVVGNTL 60
DB 1 MELAAWCRMGFLALLPPGIAGTQVCTGDMKRLRLPASPEHLDMLRLHLYGCGQVVGNTL 60
QY 61 ELTYLPTMAASLFLDIOGVGVYLIANNQVQLQRLRYRGQLPEDNALALVDNG 120
DB 61 ELTYVPAASLSFLDIOGVGVYLIANNQVQLQRLRYRGQLPEDNALALVDNR 120
QY 121 DPLNNTTPTVT-GASFGRLRLQLRLSTELIKGVLIQNNPOLCYODTILMKOIFHKNNQL 179
DB 121 DPNQVNAVSTGRTPEGLRLQLRLSTELIKGVLIQNNPOLCYODTILMKOIFHKNNQL 180
QY 180 ALTLIDTRSRACHPSCPMCKSGRCWGSSEDDQSLRTVCAGGACRCKGRLPTDCHBQ 239
DB 181 APVDIDTRSRACHPSCPMCKSGRCWGSSEDDQSLRTVCAGGACRCKGRLPTDCHBQ 240
QY 240 CAAGCTGPKHSDCLALPHNHSIGELICPALVYNTPTFESMNPBEGRTYFGASCVTAC 299
DB 241 CAAGCTGPKHSDCLALPHNHSIGELICPALVYNTPTFESMNPBEGRTYFGASCVTAC 300
QY 300 PNYLSTVGSCTVLCPLNQBVTAEADGTQRCCKSKPCARVCYGLGMEHLREVAATVSA 359
DB 301 PNYLSTVGSCTVLCPLNQBVTAEADGTQRCCKSKPCARVCYGLGMEHLREVAATVSA 360
QY 360 NIOEPAGCKKIFGSLAPLPESFDGDPASNTAPLOEOLQVETLEITGYLYISAMPDL 419
DB 361 NVQEPDCKKIFGSLAPLPESFDGDPSSGIAPLRBQLOVETLEITGYLYISAMPDL 420
QY 420 PDLSVFOULQYIRGILHNGAYSLTLOGLISWGLRLRLBELGSLALIHNTHLCPHYT 479
DB 421 PDLSVFOULQYIRGILHNGAYSLTLOGLISWGLRLRLBELGSLALIHNTHLCPHYT 480
QY 480 VPMQDLFNNPQALLHTANRPDE-CVGEGLACHOLCARGCWGPGPTQVNCOSQFLRGO 538
DB 481 VPMQDLFNNPQALLHTANRPDE-CVGEGLACHOLCARGCWGPGPTQVNCOSQFLRGO 540
QY 539 BCVEECRVLOGLPREYVNAHQLPCHPECOPOGNSVTCFGEBAADQVCAAHYKDPFCVA 598
DB 541 BCVEECRVLOGLPREYVNAHQLPCHPECOPOGNSVTCFGEBAADQVCAAHYKDPFCVA 600
QY 599 RCPGSGVKDLSYMPIMKYPDEBEGICQPCPINCSTSCVDLDRGCGPAEGRASPVTFIATV 658
DB 601 RCPGSGVKDLSYMPIMKYPDEBEGICQPCPINCSTSCVDLDRGCGPAEGRASPVTFIATV 660
QY 659 VGILLVVLGVVFGILLKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAKRILKET 718
DB 661 VGILLVVLGVVFGILLKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAKRILKET 720
QY 719 ELARKVKVLGSGAFGVYKGIWIPDGENYKIPVAILVLEBNTSPKANKELIDBAVYMAVG 778
DB 721 ELARKVKVLGSGAFGVYKGIWIPDGENYKIPVAILVLEBNTSPKANKELIDBAVYMAVG 780
QY 779 SPYLSRLIGICLTSTVOLVTOQMPYGCILLDYREKRGVLSODLLNMCQIAKGSYED 838
DB 781 SPYLSRLIGICLTSTVOLVTOQMPYGCILLDYREKRGVLSODLLNMCQIAKGSYED 840
QY 839 VRLVHRDLAARNVLYKSPNHKITDFFGLARLLIDETRYHADGGKVPKMMALSLILRR 898
DB 841 VRLVHRDLAARNVLYKSPNHKITDFFGLARLLIDETRYHADGGKVPKMMALSLILRR 900
QY 899 FTHOSDVMSYGVTWELMTFGAKPYDGI-PAREIPDLLEKGERLPQPICTIDVYVIMWKC 958
DB 901 FTHOSDVMSYGVTWELMTFGAKPYDGI-PAREIPDLLEKGERLPQPICTIDVYVIMWKC 960
QY 959 WMIDSECCPRRELVSERKARADQQRVYVIONEDLGPASPLDSTFTYSLLLEDDMGVLV 1018
DB 961 WMIDSECCPRRELVSERKARADQQRVYVIONEDLGPASPLDSTFTYSLLLEDDMGVLV 1020

```

```

QY 1019 DABEYLVEQOGFPCPDPAAGWVHNRHNSSTRSGGGDLITGLSPSEBEPAPLAPS 1078
DB 1021 DABEYLVEQOGFPCPDPAAGWVHNRHNSSTRSGGGDLITGLSPSEBEPAPLAPS 1080
QY 1079 EGAGSDVFDGDLGMGAKAGLSLPTHDPSTLQRYSEDPVTPPLSTDTGVAPLTSPOPE 1138
DB 1081 EGAGSDVFDGDLGMGAKAGLSLPTHDPSTLQRYSEDPVTPPLSTDTGVAPLTSPOPE 1140
QY 1139 VYNQDVARPOPPSPREGLPAPAPGATLERPKTSLPKGNQVVKDVFAPGAVENPEYLT 1198
DB 1141 VYNQDVARPOPPSPREGLPAPAPGATLERPKTSLPKGNQVVKDVFAPGAVENPEYLT 1200
QY 1199 PQGGAPOPHPPPPAPSPADNLVYWDQDPPEKGAPEPTFGKPTAENPEYLGLDVAV 1255
DB 1201 PREGTASPPHSPAPSPADNLVYWDQDSSQGGPPSPFEGNPTAENPEYLGLDVAV 1257

RESULT 5
ERBB2_MESAU STANDARD; PRT; 1254 AA.
AC 060553;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Receptor tyrosine-protein kinase erbB-2 precursor (BC 2.7.1.112)
DE (p185erbB2) (C-erbB-2) (NEU proto-oncogene).
GN Name=ERBB2; Synonyms=NEU;
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Eulipotyphla; Glirres; Rodentia; Sciurognathi;
OC Muridae; Cricetidae; Cricetinae; Mesocricetus.
OC NCBI_TaxId=10036;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Nerve;
RX MEDLINE=94193007; PubMed=7908275; DOI=10.1016/0378-1119(94)90553-3;
RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M., Yamazaki Y.,
RA Ishikawa T.;
RT "Cloning and activation of the Syrian hamster neu proto-oncogene.";
RL Gene 140:251-255(1994).
CC -! FUNCTION: Essential component of a neurotrophin-receptor complex,
CC although neurotrophins do not interact with it alone. Grp30 is a
CC potential ligand for this receptor. Not activated by EGF, TGF-
CC alpha and amphiregulin (By similarity).
CC -! CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + a protein
CC tyrosine phosphate.
CC -! SUBUNIT: Heterodimer with each of the other ERBB receptors
CC (Potential). Part of a complex with EGFR and either PIK3C2A or
CC PIK3C2B. Interacts with PRKABP and PLK4BL. May interact with
CC PIK3C2B when phosphorylated on Tyr-1196 (By similarity).
CC -! SUBCELLULAR LOCATION: Type I membrane protein.
CC -! PTM: Ligand-binding increases phosphorylation on tyrosine
CC residues.
CC -! SIMILARITY: Belongs to the Tyr protein kinase family. EGF receptor
CC subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL, D16295; BAA03801.1; -, mRNA.
CC PIR, I48161; I48161.
CC HSSP, P06494; INB9.
CC SMR, Q60553; 23-629.
CC InterPro: IPR000494; EGFR_L.
CC InterPro: IPR006211; Furin-like.
CC InterPro: IPR006212; Furin repeat.
CC InterPro: IPR007119; Prot_Kinase.
CC InterPro: IPR001245; Tyr_Kinase.
CC InterPro: IPR008266; Tyr_kinase_AS.
CC InterPro: IPR004019; YLP_motif.

```

	DR	Pfam; PF00757; Furin-like; 1.	
	DR	Pfam; PF01030; Recep_L_domain; 2.	
	DR	Pfam; PF02757; YLP; 2.	
	DR	PRINTS; PR00109; TYRKINASE.	
	DR	Prodom; PD000001; Prot_kinase; 1.	
	DR	SMART; SMO0261; FU; 4.	
	DR	SMART; SMO0261; Tyrc; 1.	
	DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
	DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.	
	DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.	
	KW	ATP-binding; Disease mutation; Glycoprotein; Kinase; Multigene family;	
	KV	Nucleotide-binding; Phosphorylation; Proto-oncogene; Receptor; Signal;	
	KM	Transferase; Transmembrane; Tyrosine-protein kinase.	
	FT	SIGNAL	1 21
	FT	CHAIN	22 1254
	FT	TPOD_DOM	22 652
	FT	TRANSMEM	653 675
	FT	TPOD_DOM	676 1254
	FT	DONAI_N	720 987
	FT	NP_BIND	726 734
	FT	REGION	1195 1197
	FT	COMBIAS	158 368
	FT	COMBIAS	472 644
	FT	ACT_SITE	845 845
	FT	BINDING	753 753
	FT	MOD_RES	1139 1139
	FT	MOD_RES	1196 1196
	FT	MOD_RES	1247 1247
	FT	CARBOHYD	68 68
	FT	CARBOHYD	125 125
	FT	CARBOHYD	187 187
	FT	CARBOHYD	259 259
	FT	CARBOHYD	530 530
	FT	CARBOHYD	571 571
	FT	CARBOHYD	629 629
	FT	DISULFID	195 204
	FT	DISULFID	199 212
	FT	DISULFID	236 244
	FT	DISULFID	240 252
	FT	DISULFID	255 264
	FT	DISULFID	268 295
	FT	DISULFID	299 311
	FT	DISULFID	315 331
	FT	DISULFID	334 338
	FT	DISULFID	511 520
	FT	DISULFID	515 528
	FT	DISULFID	531 540
	FT	DISULFID	544 560
	FT	DISULFID	563 576
	FT	DISULFID	567 584
	FT	DISULFID	587 596
	FT	DISULFID	600 623
	FT	DISULFID	626 634
	FT	DISULFID	630 642
	FT	VARIANT	658 658
	FT	VARIANT	659 659
	SO	SEQUENCE	1254 AA; 138253 MW; 974CG3791C21F2B81 CRC64;
		Query Match	87.9%; Score 5993.5; DB 1; Length 1254;
		Beat Local Similarity	87.7%; Pred. No. 1.le-296;
		Matches 1101; Conservative	57; Mismatches 96; Indels 1; Gaps 1
Oy		1 MELAALCWGLIALALLPPGAASTOVCTGTDMKLRIPASGETHLMRLHYOGCYOYGNTL	60
Db		1 MELAAWCWGILIALLLSPAGSGTQVCTGTDMKLRIPASGETHLDIVRHLYOCGYOYGNTL	60
Oy		61 ELTYVPTNASTLSFLDDIOEVGGYVLIAHNQGVQLQRIRIYRGTOLEFDNALAVLNG	120
Db		61 ELTYVPANATSLFLDDIQGVGYMLIASQVHVPLQRIRIYRGTOLEFDKRALVLDNR	120
Oy		121 DPLNNTTPTVGASPGGLRELQARSITELIKGVALIQRNPOLCYODTIIMKDIFHKNOQA	180

Db	121	DPDQVTTATRTTBGRLREQLRSLTEILKSGVILIRGNPOLCYDDYIMKQVFNKQOLA	180
Qy	181	LTLIDYNSRACHPCSPWCKSGRCSGESSDDQSLRTRVCAGGACRCKGRLPTDCHEQC	240
Db	181	PVDIDYNSRACPCPCAPACKONHCWGSBEDCQTLTGTAIRAVAPARARLPTDCHEQC	240
Qy	241	AAGCTGPRHNSCLACLPHNSGIELHCPALVYNTDTFSBMPNPEGRYTFGASCYTAAC	300
Db	241	AAGCTGPRHNSCLACLPHNSGIELHCPALVYNTDTFSBMPNPEGRYTFGASCYTTCP	300
Qy	301	YNYLSTDVSGCTVCPILHNOEVTADGFORCEKSKPCARVCYGLGMEHLREVAVTSAN	360
Db	301	YNYLSTDVSGCTVCPILHNOEVTADGFORCEKSKPCARVCYGLGMEHLREVAVTSAN	360
Qy	361	IOEPFAGCKKIKGSLAFIPESFSDGPNASNTAPLOEOLQVPELTBEITGYLISAMPDLP	420
Db	361	IOEPFAGCKKIKGSLAFIPESFSDGPNSSGIANLPTREQLOVPELTBEITGYLISAMPDLSH	420
Qy	421	DLSPFQNIQVIRGSIILHNGAVSLTLQGLGISWGLRSLRELGSLALIHNTHLCPVHTV	480
Db	421	DLSPFQNIQVIRGSHVLDHGAYSIALQGLGIRMLGIRSLRELGSSGLVIHRTHLCPVHTV	480
Qy	481	PWDLQFRPHOALLHTANRPEDECVSGSLACHOICARGHCHMGCPPTOCVNSOPLROEC	540
Db	481	PWDLQFRPHOALLHSGPSEEBEGLDOPALPYLCAGHCHMGCPPTOCVNSHFLRQBC	540
Qy	541	VEBCRVLOGLREYEVNARHCLPCHPECOPONGSYTCFGEPEADQCVAACHYKDPFCVARC	600
Db	541	VKECRVWKGLPREYVNGKHCLPCHPECOPONSTETCTGSEADQCTACPHYKDSFCVARC	600
Qy	601	PSGVKPDLSYMPWKFPDEBGACQPCPINCTHSCVDDLKGCPCAPQBPASPLTISAVVG	660
Db	601	PSGVKPDLSYMPWKYDDEBGMCPCEPINCTHSCVDDLBERGCPBPQBPASPLTISAVVG	660
Qy	661	ILVYVYLVGVNPGILLIKRQOKIRXYTMRILLQETELVEPLTPSGAMNQOMRLIKETEL	720
Db	661	ILFLFLVGVNPGILLIKRRQKIRXYTMRILLQETELVEPLTPSGAMNQOMRLIKETEL	720
Qy	721	RKYVULSGAGFVYKGIWIPDGENYVKI PVALKYLRENTS PKANKELILDEAYVWAGVGP	780
Db	721	RKYVULSGAGFVYKGIWIPDGENYVKI PVALKYLRENTSPKANKELILDEAYVWAGVGP	780
Qy	781	YVSRILGICLTSTVOLVTQMLPVGCLLDHVENRGRIGSODLLNWCQIAKMSYLEDVR	840
Db	781	YVSRILGICLTSTVOLVTQMLPVGCLLDHVEHRGRIGSODLLNWCQVIAKMSYLEDVR	840
Qy	841	LVHMDLAARNLVKSPNHVKITDGLARLLDIDETEVHADGGKPIKIMALESTLRRRFT	900
Db	841	LVHMDLAARNLVKSPNHVKITDGLARLLDIDETEVHADGGKPIKIMALESTLRRRFT	900
Qy	901	HOSDWSYGVVMEMLMTFGAKPYVIGIPARELIPDLJEKGERLPORPICI DIVYIMVNCMM	960
Db	901	HOSDWSYGVVMEMLMTFGAKPYVIGIPARELIPDLJEKGERLPORPICI DIVYIMVNCMM	960
Qy	961	IDSECRPRFRELVSFBSMARMDPQRFVYIQNEBGLGSPASPLDSTFYRSLLEBDDMDGLVDA	1020
Db	961	IDSECRPRFRELVSFBSMARMDPQRFVYIQNEBGLGSPASPLDSTFYRSLLEBDDMDGLVDA	1020
Qy	1021	EELYVPOQGFCCDPAPQAGCMVHRRRSSTSGGSDLTGLGEPSEEBAPRSPLASSEG	1080
Db	1021	EELYVPOQGFCCDPAPQAGCMVHRRRSSTSGGSDLTGLGEPSEEBAPRSPLASSEG	1080
Qy	1081	AGSVUFPOSDGMAKAGLQSLPHTHDBPRLORYSDDPVLVPSBNDGVVAPLTCGPOEYV	1140
Db	1081	AGSVUFBEELGMAKTKBPQISPRDLSPLQKQYSHDPTLLPLETEIDGVVAPLACSGPOEYV	1140
Qy	1141	NQOPVREPOPBPBREGPLPAARPAATATERPXTLSPGKNGVVKOVFAFGAVENPEYLTPO	1200
Db	1141	NQPEVRRQPLTPBGPRLPVRRPAGATLIERPXTLSPGKNGVVKOVFTTGAVENPEYLVPR	1200
Qy	1201	GGAAPQHPHPAPFAPFADNIYYMDQDPERGAPSTFKGPTAENPEYLGIDVFFV	1255

Db 1201 GGSASQPH-PPALCPAFDNLVYWDQDPSEBRSPPPTFEGTPTABNPEYLGLDVPV 1254

RESULT 6

Q62PE0_MOUSE PRELIMINARY; PRT; 1305 AA.

AC Q62PE0; 05-JUL-2004 (TReMBLrel. 27, Created)

DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)

DE MKIAA3023 protein (fragment).

GN Name=Erbb2; Synonym=MKIAA3023;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; OC Muridae; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RA NUCLEOTIDE SEQUENCE.

RP TISSUE=Embryonic tail;

RX MEDLINE=22977043; PubMed=14621295;

RA Okazaki N., Kikuno R., Ohara K., Inamoto S., Koseki H., Hirooka S., Saga Y., Nagase T., Ohara O., Koga H.;

RT "Prediction of the coding sequences of mouse homologues of KIAA gene: RT Ii. the complete nucleotide sequences of 500 mouse KIAA-homologous RT cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries."

RL DNA Ref. 10:167-180(2003).

DR EMBL; AK129487; BAC98297.1; -; mRNA.

DR HSSP; P13362; IAGM.

DR SNR; Q62PE0; 72-679.

DR Ensembl; ENSMUSG0000062312; Mus musculus.

DR MGI; MGI:95410; Erbb2.

DR GO; GO:0016324; C:apical plasma membrane; IDA.

DR GO; GO:0005737; C:cytoplasm; IDA.

DR GO; GO:0005515; F:protein binding; IPI.

DR GO; GO:0007507; P:heart development; IMP.

DR GO; GO:0042552; P:myelination; IMP.

DR GO; GO:0007422; P:peripheral nervous system development; IMP.

DR InterPro; IPR000494; EGFR_L.

DR InterPro; IPR006211; Furin-like.

DR InterPro; IPR006212; Furin repeat.

DR InterPro; IPR000719; Prot. Kinase.

DR InterPro; IPR002290; Ser Thr. Kinase.

DR InterPro; IPR001245; Tyr_kinase.

DR InterPro; IPR008266; Tyr_kinase_AS.

DR InterPro; IPR004019; YLP motif.

DR Pfam; PF00757; Furin-like; 1.

DR Pfam; PF01030; Recep_L_domain; 2.

DR Pfam; PF02757; YLP_3.

DR PRINTS; PR00109; TYRKINASE.

DR ProDom; PD000001; Prot_Kinase; 1.

DR SMART; SMO0261; FU; 4.

DR SMART; SMO0220; S_TKc; 1.

DR SMART; SMO0219; Tyrc; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

DR ATP-binding; Kinase; Nucleotide-binding; transferase.

FT NON TER 1 1

SEQUENCE 1305 AA; 143508 MW; A5ID897408521860 CRC64;

Query Match 87.8%; Score 5982.5; DB 2; Length 1305;

Best Local Similarity 87.7%; Pred. No. 4e-296;

Matches 1101; Conservative 55; Mismatches 99; Indels 1; Gaps 1;

Qy 1 MELALACRWGLLLLPAGASTQVCTGDMKRLRLPASPETHLMDLRLYGGCOVVGNTL 60

Db 50 MELAAKRWGLLLLSFGAAGTQVCTGDMKRLRLPASPETHLMDLRLYGGCOVVGNTL 109

Qy 61 ELTYLPTNASLSTFDIOGVGVYLIANNOVROVPLQRLRIVRGTQLFEDNYALAVLDNG 120

Db 110 ELTYLPTNASLSTFDIOGVGVYLIANNOVROVPLQRLRIVRGTQLFEDNYALAVLDNR 169

Qy 121 DPLANN-TPYVAGSPGRLREQLRSITLILKGVLIQNPNQCYODITLMDIIFKNNQL 179

Db 170 DPLDNTTAAARGRTPBEGRLQLRSITLILKGVLIQNPNQCYODITLMDIIFKNNQL 229

Qy 180 ALTLIDTNRSAPACHPSPWCKSRGCRGSSSDCQLRTVACAGCARKGLPTDCCHQ 239

Db 220 APVMDITNRSAPACHPCAPTCQDNHCWGBSPBDCLITGTICTSGCARKGLPTDCCHQ 289

Qy 240 CAAGCTGPKSDCLACLPNNSGICELHCPALVYNTDTPESMPNBERYFGASCYAC 299

Db 290 CAAGCTGPKSDCLACLPNNSGICELHCPALVYNTDTPESMPNBERYFGASCYAC 349

Qy 300 PNYLSTVGSCTVCPPLANOEYTAEDTQRCCKSKPCARVCYGLMEHLREYAVTSA 359

Db 350 PNYLSTVGSCTVCPPLANOEYTAEDTQRCCKSKPCARVCYGLMEHLREYAVTSA 409

Qy 360 NIOEPAGCKKI FGSIAFLPESFDGDPASNTAPLQPEOLQVEFTLEITGYLYISAMPDSL 419

Db 410 NIOEPAGCKKI FGSIAFLPESFDGDPASNTAPLQPEOLQVEFTLEITGYLYISAMPDSL 469

Qy 420 PDLSPQNLQYIRGRILNNGAYSLTQGLDLSWGLSRLBSGLALIHNTILCFVNT 479

Db 470 QDLSPQNLQYIRGRILNNGAYSLTQGLDLSWGLSRLBSGLALIHNTILCFVNT 529

Qy 480 VPMQDLFNPQHALLHTANRPEDECVGGLACHOGLCARGHCWGPPTQCVNCSQFLRGQ 539

Db 530 VPMQDLFNPQHALLHTANRPEDECVGGLACHOGLCARGHCWGPPTQCVNCSQFLRGQ 589

Qy 540 CVBECRVLOGLPREYVNAHCLPCHPECOFNGSVTCFGEPAQCVACAHYKDPFCYAR 599

Db 590 CVBECRVLOGLPREYVNAHCLPCHPECOFNGSVTCFGEPAQCVACAHYKDPFCYAR 649

Qy 600 CPSCVKPDLSTMPYWKPPDEEGACQPCINCTHSCVLDLMDGCPAEOASLTSISAVV 659

Db 650 CPSCVKPDLSTMPYWKPPDEEGACQPCINCTHSCVLDLMDGCPAEOASLTSISAVV 709

Qy 660 GILLVVLVGVVFGILIRKROOKIRKTYMRRLLOETELVEPILTPSGAMPNQOMRLKETE 719

Db 710 GILLVVLVGVVFGILIRKROOKIRKTYMRRLLOETELVEPILTPSGAMPNQOMRLKETE 769

Qy 720 LRKVKVLSGAFGVYVYKGIWIPDGENYKIPVAILVLRBNTSPKANKEILDEAYVAVGVS 779

Db 770 LRKVKVLSGAFGVYVYKGIWIPDGENYKIPVAILVLRBNTSPKANKEILDEAYVAVGVS 829

Qy 780 PYVSRLLGICLTSTVQLVTLQMPYGCILDHTRKRGRLGSDLLNMCQIAGSYLEDV 839

Db 830 PYVSRLLGICLTSTVQLVTLQMPYGCILDHTRKRGRLGSDLLNMCQIAGSYLEDV 889

Qy 840 RLVRHDLAARVLYKSPNHVKITDGLARLIDIDETEHYADGGKVPIMMALLESILRRRF 899

Db 890 RLVRHDLAARVLYKSPNHVKITDGLARLIDIDETEHYADGGKVPIMMALLESILRRRF 949

Qy 900 THQSDVMSYGVTVWELMTFGAKPYDGI PARAI PDLLEKGERLPPOPCTIDVYIMVWCW 959

Db 950 THQSDVMSYGVTVWELMTFGAKPYDGI PARAI PDLLEKGERLPPOPCTIDVYIMVWCW 1009

Qy 960 MIDSECRPRFRELVSERSMARDPQRFVYIYNEDLGPASPLDSTFTRSLLEDDMDGLVD 1019

Db 1010 MIDSECRPRFRELVSERSMARDPQRFVYIYNEDLGPASPLDSTFTRSLLEDDMDGLVD 1069

Qy 1020 AAEYLVPOQGFPCDDPA PGAGCWVHHHRH86STSGGGDLTLGLPSESEEARPLAPSE 1079

Db 1070 AAEYLVPOQGFPCDDPA PGAGCWVHHHRH86STSGGGDLTLGLPSESEEARPLAPSE 1129

Qy 1080 GAGSDVFDGDLGMAAGKAGSLPTHDSPPLQRYSEDPVPLPSESTDGVAVLPTCSPOPEY 1139

Db 1130 GAGSDVFDGDLGMAAGKAGSLPTHDSPPLQRYSEDPVPLPSESTDGVAVLPTCSPOPEY 1189

Qy 1140 VNQEDVRRPQPSPRBGLPAARPGATLBERKXTSPGKNGYVYKQVAFAGAVENPEYITP 1199

Db 1190 VNQEDVRRPQPSPRBGLPAARPGATLBERKXTSPGKNGYVYKQVAFAGAVENPEYITP 1249

Qy 1200 OGGAAPPHPPAPSPAFDNLVYWDQDPPEKGAPESTFKGTPTAENBEYLGLDVPV 1255
 Db 1250 RAGTASGPHSPAPSPAFDNLVYWDQDSSEGGPPSTFEETPTAENBEYLGLDVPV 1305

RESULT 7
 ID Q8COE7 MOUSE PRELIMINARY; PRT; 881 AA.
 AC Q8COE7;
 DT 01-MAR-2003 (TrEMBL:rel. 23, Created)
 DT 01-MAR-2003 (TrEMBL:rel. 26, Last sequence update)
 DE Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
 DE enriched library, clone:603049F08 product:v-erb-b2 erythroidlastic
 DE leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene
 DE homolog (avian), full insert sequence. (Fragment).
 GN Name=ErbB2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:119-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikita I., Pesole G., Quackenbush J.,
 RA Schirini L.M., Saudill F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barth G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Ranciocich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
 RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA The PANTOM Consortium,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;

RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsuno H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitt. (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK031542; BAC27442.1; -; mRNA.
 DR HSSP; P06494; INBY.
 DR SMR; Q8COE7; 1-255.
 DR Ensembl; ENSMUSG00000062312; Mus musculus.
 DR MGI; MGI:95410; Etbd2.
 DR GO; GO:0016324; C:apical plasma membrane; IDA.
 DR GO; GO:0005737; C:cytosol; IDA.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:0007507; P:heart development; IMP.
 DR GO; GO:0042552; P:myelination; IMP.
 DR GO; GO:0007422; P:peripheral nervous system development; IMP.
 DR InterPro; IPR000494; EGFR_L.
 DR InterPro; IPR006212; Furin_repeat.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR InterPro; IPR008266; Tyr_kinase_AS.
 DR InterPro; IPR004019; YLP_motif.
 DR Pfam; PF01030; Recep_L_domain; 1.
 DR Pfam; PF02757; YLP; 2.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00261; FU; 2.
 DR SMART; SM00219; Tyrc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 KW ATP-binding; Kinase; Nucleotide-binding; Transferase;
 KW Tyrosine-protein kinase.
 FT NON_TER 1
 SQ SEQUENCE 881 AA; 97501 MW; 5D5042BE9F80836 CRC64;

Query Match 61.9%; Score 4216; DB 2; Length 881;
 Best Local Similarity 88.3%; Pred. No. 2 2e-206;
 Matches 778; Conservative 39; Mismatches 64; Indels 0; Gaps 0;

Qy 375 AFLPESFDGDPASNTAPLOEOLQVEFTLEITGYLYISAMPDLSPLDLSYFONLQVIRGR 434
 Db 1 AFLPESFDGDPASNTAPLOEOLQVEFTLEITGYLYISAMPDLSPLDLSYFONLQVIRGR 60

Qy 435 ILNNGAYSLTGLGLSWLGLRLRLSGSLALIHNTHTCFPHYTPWMDLPNNPQALL 494
 Db 61 ILNNGAYSLTGLGLSHSLRLRLDGLALIHNTHTCFPHYTPWMDLPNNPQALL 120

Qy 495 HTANRPEDECVGSLACHOLCARGHCWGPPTGCNCSCQPLRGOECVBEQVAGLPREY 554
 Db 121 HSGNREBACGLBGLVNCNSLCARGHCWGPPTGCNCSCQPLRGOECVBEQVAGLPREY 180

QY	555	UNABRCLPCHRESCOPONGSVTCFGRPADOCVA CANXKDPFCFVAPRPSGVKXDSLMPYTW	614
Db	191	VGKIKCLPCHPBCQPNSSSETCYGSEADCCCAHKKSSSCVAPRPSGVKXDSLMPYTW	240
QY	615	KPDEEGACQPPICNTHS CVLDLDGCAEORASPLTSGIIISAVVGIILVVVLGVVFGIL	674
Db	241	KYPDEEGICQPPICNTHS CVLDLDGCGAEOBASVYTTI IATVGVLLFLIIVVIGIL	300
QY	675	IKRROQKIRKYTMRRLLQETELVEPLTPSGAMPNOAMRILKETELRKVVLGSAFGTV	734
Db	301	IKRRROKIRKYTMRRLLQETELVEPLTPSGAVPNQAMRILKETELRKVLGSAFGTV	360
QY	735	YKGVIIPROGENKIPAIKVLAEKNSPKANKSILDEAYMAVGSPYVSRLIGICTSTV	794
Db	361	YKGVIIPROGENKIPAIKVLAEKNSPKANKSILDEAYMAVGSPYVSRLIGICTSTV	420
QY	795	QLVLTQMPGCLLDHYRENRGRIGSODLLNMCQIAGKSYLEDVLYVRDLAARVLYK	854
Db	421	QLVLTQMPGCLLDHYRENRGRIGSODLLNMCQIAGKSYLEDVLYVRDLAARVLYK	480
QY	855	SPNHVKITDFGLARLLDIDETEVHADGKVPIMKMALESILRRPFTQSDVMSYGVYWE	914
Db	481	SPNHVKINDFGIARLLDIDETEVHADGKVPIMKMALESILRRPFTQSDVMSYGVYWE	540
QY	915	LMTFGAKPRPDGIPAEIPIPLKKGRLPPRPCTIDVYIMYKCMWIDSECRPRELVS	974
Db	541	LMTFGAKPRPDGIPAEIPIPLKKGRLPPRPCTIDVYIMYKCMWIDSECRPRELVS	600
QY	975	EFSSRARDPQRVVVIONEDLGPASPLDSFTFYRSILEDMDGDLVPAEYLVVQOGFCFD	1033
Db	601	EFSSRARDPQRVVVIONEDLGPSSPMDSTFYRSILEDMDGDLVPAEYLVVQOGFCFD	660
QY	1035	PAPGAGWVHNHRSSSTRSGGCDLTLGLEPSEBEAPRSLPAPSBGAGSDVDFDGLGMGA	1093
Db	661	PALGTGSTAHRHRSSASAGGGLTLGLEPSEBEPRRSLPAPSBGAGSDVDFDGLAVGV	720
QY	1095	AKGLQSLPTHDPSPIQRYSEDPVPLPSTDGTGVAPLTCSPOPEYVNOQVPRPOPSPRE	1155
Db	721	TKGLQSLSPHNDSPIQRYSEDPVPLPPEPTDGVVAPLACSPQPEYVNOQVPRPOPSTLPE	780
QY	1155	GPLPAPAPAGATLEPRKTLSPCKNGVVKQVAFPGAVENPEVLTQGGAAPOPHPPAPS	1211
Db	781	GPPPIPRPAGATLEPRKTLSPCKNGVVKQVAFPGAVENPEVLTAPAGASQPHSPAPS	840
QY	1215	PAFDNLYYWDODPPERGAPPTSTFKGTPYLAENBEYLGLDVVPV	1255
Db	841	PAFDNLYYWDQNSSEQPPSTFEGTPTLAENBEYLGLDVVPV	881
RESULT 8			
QSEBY4 BRARE			
ID	QSEBY4	BRARE PRELIMINARY,	PRT, 1275 AA.
AC	QSEBY4	10-MAY-2005 (TREMBlrel. 30, Created)	
DT	10-MAY-2005	(TREMBlrel. 30, Last sequence update)	
DT	10-MAY-2005	(TREMBlrel. 30, Last annotation update)	
DE	Brb2.		
GN	Name=erbb2; ORFNames=zgc:63601;		
OS	Brachydanio rerio (Zebrafish). (Danio rerio).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;		
OC	Cyprinidae; Danio.		
OK	NCBI_Taxid=7955;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	PubMed=15797019;		
RA	Lyons D.A., Pogoda H.M., Voas M.G., Woods I.G., Diamond B., Nix R.,		
RA	Arata N., Jacobs J., Talbot W.S.;		
RT	"erbb3 and erbb2 Are Essential for Schwann Cell Migration and		
RL	Myelination in Zebrafish."		
RL	Curr. Biol. 15:513-524(2005).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE.		

RA Pogoda H.-M., Voas M.G., Talbot M.S.:
 "Analysis of a zebrafish erbb2 gene."
 RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases
 DR EMBL, AY14799; AAX09341.1; -, mRNA.
 DR ZFIN, ZDB-GENE-031118-121; zgc:63601.
 DR InterPro: IPRO00494; EGFR_L.
 DR InterPro: IPRO06211; Furin-like.
 DR InterPro: IPRO00719; Furin repeat.
 DR InterPro: IPRO02290; Ser_thr_kinase.
 DR InterPro: IPRO01245; Tyr_kinase.
 DR InterPro: IPRO08266; Tyr_kinase_AS.
 DR Pfam, PF00757; Furin-like; 1.
 DR Pfam, PF00069; Kinase; 1.
 DR Pfam, PF01030; Recep_L domain; 2.
 DR PRINTS, PR00109; TYRKINASE.
 DR ProDom, PD000001; Prot_kinase; 1.
 DR SMART, SM00261; FU; 3.
 DR SMART, SM00220; S_TKC; 1.
 DR SMART, SM00219; TYRKc; 1.
 DR PROSITE, PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE, PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE, PS00109; PROTEIN_KINASE_TYR; 1.
 SQ SEQUENCE 1275 AA; 14082 MW; 67462B4FLTB6492P CXC64

Query Match	51.6%;	Score 3514.5;	DB 2;	Length 1275;
Best Local Similarity	54.9%;	Pred. No. 1.8e-170;		
Matches 721;	Conservative 148;	Mismatches 315;	Indels 129;	Gaps 28

QY	9	AGLLIALLP	PGAA	SV	OV	CT	GM	KRL	PAS	PETH	LD	ML	SH	LV	OG	SV	OV	GN	LE	TY	PT	68
Db	11	VWLL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	LL	
QY	69	ASLS	FL	OD	IQ	EV	OG	YVL	IA	HO	VO	PL	QR	LR	I	VR	GT	OL	P	ED	NY	
Db	71	PL	SL	FL	Q	IE	VE	OG	YVL	IA	HS	VS	RL	PD	NR	LL	IR	SO	L	Y	KS	
QY	129	VT	GA	S	PG	LR	EL	QL	AS	TR	EL	IL	KG	VL	IQ	NR	PO	L	Y	OT	LD	
Db	124	-	SS	OG	L	Q	IR	EL	RL	RL	SL	TE	IL	LG	YV	I	WN	P	OL	C	FR	
QY	189	SR	AC	HP	C	SP	MG	K	SR	CM	GS	ES	ED	CS	L	TR	YV	AG	CA	R	C	
Db	193	N	-	CR	CS	AC	K	S	G	G	C	M	E	K	OD	C	OT	L	S	V	N	
QY	249	HS	D	CL	AC	L	H	FN	HS	G	I	CEL	H	CP	AL	V	T	Y	N	T	D	
Db	241	D	S	D	CL	AC	H	F	N	D	S	G	C	R	P	T	I	Y	D	P	I	
QY	309	GS	CT	IV	C	EL	H	Q	EV	TA	-	ED	G	-	T	OR	C	E	K	S	R	
Db	301	-	AC	T	W	C	P	R	AN	K	EV	S	V	E	P	D	G	Q	ET	K	C	
QY	365	AG	CK	I	P	GS	L	A	P	E	S	F	D	D	P	A	S	T	A	P	L	
Db	360	T	C	E	K	I	Y	G	S	L	A	P	S	S	F	K	N	A	D	P	T	
QY	425	F	ON	I	V	I	R	IL	H	N	G	A	Y	S	L	T	O					

```

Qy 663 LVVAVGVFGLIKRRQOKIRKRYTMRLLQETELVEPTSPGAMPNOMILKETELRK 722
Dy 656 LFIILLALVFLYLRQKQKQKKEKIRRLQHELEVEPTSPGAMPNOMKILKETELIK 715
Qy 723 VKVUGAGFVTVVGIWTPDEENKIKPIVATKVRKENTSPKANKELIDBAYMAVGSPYV 782
Dy 716 LRVVSGAFVGVFVGIWTPDEENKIKPIVATKVRKENTSPKANKELIDBAYMAVGSPYV 775
Qy 783 SRLIGICLTSTVQVLTQAMPYGCCLLDHYENRGRLSGODLLMCMQIAKMSYLEDEVLY 842
Dy 776 CRLIGICLTSTVQVLTQAMPYGCCLLDHYENRGRLSGODLLMCMQIAKMSYLEDEVLY 835
Qy 843 HRDIAARNVLYKSPNHVKITDFGLARLIDDETEYHADGKVPYIKMALESILRRRTHQ 902
Dy 836 HRDIAARNVLYKSPNHVKITDFGLARLIDDETEYHADGKVPYIKMALESILRRRTHQ 895
Qy 903 SDVWSYGVTVWELMTFGAKPYESFOARDIPBLLEAGEBLSOPCNCTKEVYIMVYKCMQID 962
Dy 896 SDVWSYGVTVWELMTFGAKPYESFOARDIPBLLEAGEBLSOPCNCTKEVYIMVYKCMQID 955
Qy 963 SECPREFELVSEFSRMARDPQRFVVIQNEP-LGPASPLDSTFYRSLT--EDDMGDLVD 1019
Dy 956 PDNRPRFDVDEFTTMAKRDSPRYVIVIONEDOMLSSTVDSFPRILMAEGGVKERTD 1015
Qy 1020 AEEYLVPOQG--FFCPDPAPAGAGVHHRHSSSTR-----SGGGDI----- 1059
Dy 1016 AEEYLVPOQGISIFMTHGEMRANSPSRHSHRSTQMVVEVDGLPNGRELYSSVSMIEOSQY 1075
Qy 1060 -TL-----GLESEERAP--RSPPLASBEGASD-VF-DDDLGMGAAGKGLPTHDPS 1107
Dy 1076 PTLVVGATANMGWGTQYPPARLSISHRSAGQSDVFLDYVE-----DSCP--PS 1125
Qy 1108 PLOGYSEDPVPL-----PSETTG-----VYAPLTCSFPQRYVYVQPDVPPQSPRREGPL 1157
Dy 1126 SPCHYSKDPMPNNGIDGLDLETDMNVFLSHLPRGTHQPEYVQ----- 1170
Qy 1158 PAAPDAGATLERPKTLPSPK-----NGVKCVFAFGAVENDEYITPQGGAPQPH 1209
Dy 1171 -DMSASRSPSTL--PRKASERRFILNGL-----STGNSVENPEYVLPIGSIITPT-- 1215
Qy 1210 PPAASPAADNLVYMDQDPEERGAPSPSTPK-----TPFAENVEYIGL 1251
Dy 1216 ---SPAFDNPHYHDIAAKAQAVARVALINGTNRHQPNGFMTPFAENVEYIGL 1264

RESULT 9
Q80Y89 MOUSE PRELIMINARY; PRT; 711 AA.
ID Q80Y89_MOUSE PRELIMINARY; PRT; 711 AA.
AC Q80Y89;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE V-erb-b2 erythroblastic leukemia viral oncogene homolog 2,
neuro/10blablastoma derived oncogene homolog, isoform 2.
GN Name=Erbb2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603999;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Dapchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepien M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Borek S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Kiryushin N.I., Skalska U., Smalins D.E.,
RA Smerchek A., Schein J.E., Jones S.J.M., Maira M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046811; AAH46811.1; -; mRNA.
DR EMBL; BC053078; AAH53078.1; -; mRNA.
DR HSSP; P06494; INBX.
DR SMK; Q80Y89; 23-630.
DR Ensemble; ENSMUSG0000062312; Mus musculus.
DR MGI; MGI:95410; Erbb2.
DR GO; GO:0016324; C:apical plasma membrane; IDA.
DR GO; GO:0005515; P:protein binding; IPI.
DR GO; GO:0007507; P:heart development; IMP.
DR GO; GO:0042552; P:myelination; IMP.
DR GO; GO:0007422; P:peripheral nervous system development; IMP.
DR InterPro; IPR00494; EGFR_L.
DR InterPro; IPR06212; Purin repeat.
DR Pfam; Pf00757; Furin-like; 1.
DR Pfam; Pf01030; Recep_L_domain; 2.
DR SMART; SM00261; FU; 4.
SQ SEQUENCE 711 AA; 78707 MW; 682B18EB0E71318 CRC64;

Query Match 47.8%; Score 3256.5; DB 2; Length 711;
Best Local Similarity 84.5%; Pred. No. 1,2e-157;
Matches 589; Conservative 40; Mismatches 67; Indels 1; Gaps 1;
Qy 1 METALATCRWGLLALALPPGAASVQVCTGDMKRLPASPTHLDMLBHYGSCVVOGNTL 60
Dy 1 MELAAKCRMGFLALLSPGAAGVQVCTGDMKRLPASPTHLDMLBHYGSCVVOGNTL 60
Qy 61 ELTYLPTNASISPLQDIOEVQGVYLIHNOVROYPLQRLIRVGTQLPEDNYALAVDNG 120
Dy 61 ELTYLPANASISPLQDIOEVQGVYLIHNRVYKVPLORLIRVGTQLPEDNYALAVDNR 120
Qy 121 DPLNN-TTPYTGASPGGLARELQRLTEILKGGVLIQNRPOLCYQDNIIMWDIHKNNQL 179
Dy 121 DPLDNNVTAAAPGRPEELARELQRLTEILKGGVLIQNRPOLCYQDNIIMWDVLRKNNQL 180
Qy 180 ALTLIDNRSRACHPGSPMKGSRGSESESDCSLRTVACAGACAKGFLPTDCHEQ 239
Dy 180 APVMDNRSRACHPGSPCKDNHCKGESPEDCOILGCTTSCACAKGFLPTDCHEQ 240
Qy 240 CAAGCTGPKHSDCLAHFNHSGICELHCPALVTYNTDFESMNPBGRYTFGASCVTAC 299
Dy 241 CAAGCTGPKHSDCLAHFNHSGICELHCPALVTYNTDFESMNPBGRYTFGASCVTTC 300
Qy 300 PYNTLSTDVSGCTVYVGLAHQEVTAEGTORCEKCSRPCARVCYGLGMEHRLREYATYSA 359
Dy 301 PYNTLSTEVSSCTVYVCPNNQEVTAEDGTORCEKCSRPCGVYGLGMEHRLRGARATSD 360
Qy 360 NIOFACCKKIFGSLATLPSPSPDGPASNTAPLOPBOLOYETLEITGYLIYISAMPDSL 419
Dy 361 NIOFACCKKIFGSLATLPSPSPDGPASNTAPLOPBOLOYETLEITGYLIYISAMPESF 420

```

Oy		420	PDLSTFQNLQVIRGRILHNGANSYSLTQIGISIMLGLRSRLRELGSGLALIHNNHLCPVRT	479
Dd		421	ODLSVFQNLRVRIGRILHDGANSYSLTLQGIGHSLGLRSIRELGSGALLIHRNHLCEVNT	480
Oy		480	VPMDDLFRRPHOALLHTARPEDECGVEGLACHQLCARGHCMGPPTOCVNCQSOLFQGE	539
Dd		481	VPMDDLFRRPHOALLHSGRNPBEACGLBGLVCNSLCARGHCMPGPFQCNCQFLRGZE	540
Oy		540	CVEBCRVLIQGLEPREVYNABHCLCEBECQFONGSVTCFGFEADQCVAAHYXDPFFCVAR	599
Dd		541	CVEBCRWKMGKLPREVYRGKHCLPCBHECQFONSSEFCYSGEADQCECAHYXOSSCVAR	600
Oy		600	CPSSGVKPDLSYMPIMKFPPEBGACQCPINCTHSCVDLDKGCFAEORASPLTSISAVY	659
Dd		601	CPSSGVKPDLSYMPIMWKYPPEBGICQCPINCTHSCVDLDERGCCPAEORASPVTFIATVV	660
Oy		660	GILLVVLGVVFGLILIKRROQKRKYTMRLLOETEL	696
Dd		661	GVLLFLIIIVVIGILIKRRRKOKIRKYTMRLLOETEY	697
<hr/>				
RESULT 10				
ID	O9QX70_RAT	PRELIMINARY;	PRT; 1209 AA.	
AC	O9QX70;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Epidermal growth factor receptor.			
GN	Name=Egfr;			
OS	Rattus norvegicus (Rat).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			
CC	Muridae; Murinae; Rattus.			
OX	NCBI_TaxId=10116;			
RN				
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=Fisher; TISSUE=Liver;			
RX	MEDLINE=90258688; PubMed=2342466;			
RA	Petch L.A., Harris J., Raymond V.W., Blaband A.J., Lee D.C.,			
RA	Earp H.S.,			
RT	"A truncated, secreted form of the epidermal growth factor receptor is			
RT	encoded by an alternatively spliced transcript in normal rat tissue.";			
RL	Mol. Cell. Biol. 10:2973-2982(1990).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=Fisher; TISSUE=Liver;			
RA	Guttridge K., Dawson T.L., Earp H.S.,			
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.			
RP	[3]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=Fisher; TISSUE=Liver;			
RA	Petch L.A.,			
RL	Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; M37394; AAF14008.1; -, mRNA.			
DR	PIR; A36325; A36325.			
DR	HSSP; O9H2C9; 1M17.			
DR	SMR; O9QX70; 25-525, 27-636.			
DR	Ensembl; ENSRNOG0000004332; Rattus norvegicus.			
DR	RGD; 2543; Egfr.			
DR	GO; GO:0016020; Cmembrane; IEA.			
DR	GO; GO:0005524; P:ATP binding; IEA.			
DR	GO; GO:0005489; P:electron transporter activity; IEA.			
DR	GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.			
DR	GO; GO:0005506; F:ion ion binding; IEA.			
DR	GO; GO:0004872; F:receptor activity; IEA.			
DR	GO; GO:0006118; P:electron transport; IEA.			
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.			
DR	GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . ; IEA.			
DR	InterPro; IPR001450; 4FeS_ferredoxin.			
DR	InterPro; IPR000494; EGFR_L.			
DR	InterPro; IPR006211; Furin-like.			
DR	InterPro; IPR006212; Furin repeat			

Query Match	46.5%	Score 3172;	DB 2;	Length 1209;
Beet Local Similarity	50.2%	Pred. No. 4,6e-153;		
Matches	644;	Conservative 166;	Mismatches 348;	Indels 124;
				Gaps 27
DR InterPro; IPR000719; Prot. kinase.				
DR InterPro; IPR001245; Tyr_kinase.				
DR InterPro; IPR008266; Tyr_pkinase_AS.				
DR Pfam; PF00757; Furin-like; 1.				
DR Pfam; PF01030; Recep_L domain; 2.				
DR PRINTS; PR00353; 4F6SFRDOKIN.				
DR PRINTS; PR00109; TYRKINASE.				
DR Prodom; PD000001; Prot_kinase; 1.				
DR SMART; SMO0261; FU; 4.				
DR SMART; SMO0219; TyrcK; 1.				
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.				
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.				
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.				
KW Receptor.				
SEQUENCE	1209 AA;	134890 MW;	96FE7EF6CC1B7773 CRC64;	
Query Match	46.5%	Score 3172;	DB 2;	Length 1209;
Beet Local Similarity	50.2%	Pred. No. 4,6e-153;		
Matches	644;	Conservative 166;	Mismatches 348;	Indels 124;
				Gaps 27
DR InterPro; IPR000719; Prot. kinase.				
DR InterPro; IPR001245; Tyr_kinase.				
DR InterPro; IPR008266; Tyr_pkinase_AS.				
DR Pfam; PF00757; Furin-like; 1.				
DR Pfam; PF01030; Recep_L domain; 2.				
DR PRINTS; PR00353; 4F6SFRDOKIN.				
DR PRINTS; PR00109; TYRKINASE.				
DR Prodom; PD000001; Prot_kinase; 1.				
DR SMART; SMO0261; FU; 4.				
DR SMART; SMO0219; TyrcK; 1.				
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.				
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.				
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.				
KW Receptor.				
SEQUENCE	1209 AA;	134890 MW;	96FE7EF6CC1B7773 CRC64;	
3 LAALCRWGLLALALPPGA-ASTQVCTGTDMLRLPASPEYHLDMLRHLYQGQVYQGNLE	61			
15 LAALCAAG-----GALBEKKVQGGSNRLTGQGTEDHFLSLQRMHNCCEVILGNLE	66			
62 LTYLPTNASLSTLDIOYQGVYLIAHNOVROLRLRYNGTOLFEDNYALALVNDND	121			
67 IYYVGRNLDLSFLKTOEAVGYLLALNTAVRIPLENIQIIRGNALYENTYVALVLSN--	124			
122 PLNNTTPYTGASPGGLRELOLRSLTEILKGGYLIGNPOLCYDPTILWMDIFHKNNQLAL	181			
125 -----YGNKKTLRELPMKNLOEILIGAARFENNPLFCMMETIIQMDIV-QDVFLSN	175			
182 TLIDITNRS-RACHPESPCKSGSRGESSBDQSLTRTVCAAGCA-RCKGRPLTDCBQ	239			
176 MSMVDQRHLTGCPKDDPSCPNSSCMGRBENQKLTIIICAQGGRCKGRGSPSDCHNQ	235			
240 CAAGCTGPRHSDCIACLHPNHSIGTEILHCPALVTYNTDTFESMPNDEGRYTTGASCVTAC	299			
236 CAAGCTGPRSDCLVCHFRDEATCKDCTCPMLTLYNPPTYQMDVNPBGKYSFGATCVKXC	295			
300 PNYLTSTVGSSTLYVCPRLNQGVTAEDGTQREKSKSCARCYGLGMEHLREVRALVTA	359			
296 PNNYVYTHGSCVRAQCPDYEV-EDGYSKKKCDGPRKQCNIGIGEFKOTLSINAT	354			
360 NIOEPAGCKKIFGSLAFLPESFPDGPASNTAPLOEQLQVFETLEITGYLYISAMPDL	419			
355 NIKHKYCTAISGDHLIPVAFKGSFTRTPRLDRRELEIKTVAEITGFFLLIQAMPBW	414			
420 PDLVSFONLQVIRGRILHNGAVSLTLQGLISWLGRLSRLREISGLALIHNTHLCEVHT	479			
415 TDLHAFENLEIRGRTHQGGFSILAVVGNLITSLGRSLKEISDGDVILISGRNLCEYAT	474			
480 VPMDDLFRNPDLALHTANPREDECVESGLACHQICAGHCMGPRPTQVNCSSQFLRGE	539			
475 INMKKLFQTPNOKTKIMNRAEKDKATNHNVCNPLCSSBGCGPPTDVCSSQNVSRGE	534			
540 CYEECRVLQGLPREYVYNAARHCLPCHPECOPONGUSYCFGEPAEDQVCAACHYDPFCVVR	599			
535 CVDKCNILIEGEPREVENSECTQCPRPCLPOTMAMITTCGRGDNCKIKCAHYDGHCVYT	594			
600 CPSSGVKPLSYMPIMKFPDEBGAQCPENCTHSCVDDLKDGCPAERASP-LTSIIISAV	658			
595 CPSSGIMGNNTLL-VKFPADANNVCHLCHANCTYGGAGPRLKGC--QDPGPKPIPSIATGI	651			
659 VGIILVVVLGVVFGI-LIKRQOKIRKTYTKRLLDTELYEVLPTSSGAMPNQAQRILKE	717			
652 VGGFLFIIV-VALGIGLFMRRRQLVRKTLRRLLOERELVEPLTSSGEAPNQAHLRIKE	710			
718 TELRKVTKLGSAGFGYVKGKGIIPGGENVKIPVALIKVLRENTSPRANKIIDEAYVMAGV	777			
711 TEFKKIKVLGSGAFGVYKGLMIPGEKVKIPVALIKELREATSPRANKIIDEAYVMASV	770			
778 GSPVYSRLIGICLTSTVOLVTQLMPEYGLLDHVRNRRGLSGODILNMCQITAKGMSYLE	837			

RL Oncogene Res. 1:375-396(1987).
 RN [12]
 RP NUCLEOTIDE SEQUENCE OF 1-29.
 RX MEDLINE=91107677; PubMed=1988448;
 RA Haley J.D., Waterfield M.D.;
 RT "Contributory effects of de novo transcription and premature
 RT transcript termination in the regulation of human epidermal growth
 RT factor receptor proto-oncogene RNA synthesis.";
 RL J. Biol. Chem. 266:1746-1753(1991).
 RN [13]
 RP NUCLEOTIDE SEQUENCE OF 1-29.
 RX MEDLINE=85270438; PubMed=2991899;
 RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;
 RT "Characterization and sequence of the promoter region of the human
 RT epidermal growth factor receptor gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).
 RN [14]
 RP NUCLEOTIDE SEQUENCE OF 25-49.
 RX MEDLINE=84172183; PubMed=6324343;
 RA Weber W., Gill G.N., Spiesse J.;
 RT "Production of an epidermal growth factor receptor-related protein.";
 RL Science 224:294-297(1984).
 RN [15]
 RP PROTEIN SEQUENCE OF 540.
 RA Kohda D.;
 RL Submitted (SEP-1997) to Swiss-Prot.
 RN [16]
 RP PROTEIN SEQUENCE OF 687-705; 986-998; 1000-1023; 1026-1030 AND
 RP 1068-1077, AND PHOSPHORYLATION SITES THR-693; SER-695; SER-1070 AND
 RP SER-1071.
 RX MEDLINE=88330814; PubMed=3138233;
 RA Heisermann G.J., Gill G.N.;
 RT "Epidermal growth factor receptor threonine and serine residues
 RT phosphorylated in vivo.";
 RL J. Biol. Chem. 263:13152-13158(1988).
 RN [17]
 RP PROTEIN SEQUENCE OF 25-39.
 RX PubMed=15340161; DOI=10.1101/pa.04682504;
 RA Zhang Z., Hentzel W.J.;
 RT "Signal peptide prediction based on analysis of experimentally
 RT verified cleavage sites.";
 RL Protein Sci. 13:2819-2824(2004).
 RN [18]
 RP PROTEIN SEQUENCE OF 740-744 AND 746-747.
 RX MEDLINE=85182650; PubMed=2985580;
 RA Russo M.W., Lukac T.J., Cohen S., Staros J.V.;
 RT "Identification of residues in the nucleotide binding site of the
 RT epidermal growth factor receptor/kinase.";
 RL J. Biol. Chem. 260:5205-5208(1985).
 RN [19]
 RP RECEPTOR ACTIVITY.
 RX MEDLINE=84191554; PubMed=6325948;
 RA Micczkowiak B., Mosig G., Cohen S.;
 RT "ATP-stimulated interaction between epidermal growth factor receptor
 RT and supercoiled DNA.";
 RL Nature 309:270-273(1984).
 RN [20]
 RP LIGAND BINDING.
 RX MEDLINE=90003233; PubMed=2790960; DOI=10.1016/0092-8674(89)90867-2;
 RA Chen W.S., Lazar C.S., Lund K.A., Welsh J.B., Chang C.P., Walton G.M.,
 RA Der C.J., Wiley H.S., Gill G.N., Rosenfeld M.G.;
 RT "Functional independence of the epidermal growth factor receptor from
 RT a domain required for ligand-induced internalization and calcium
 RT regulation.";
 RL Cell 59:33-43(1989).
 RN [21]
 RP PHOSPHORYLATION.
 RX MEDLINE=89278137; PubMed=2543678;
 RA Margolis B.L., Lax I., Kriz R., Dombalagian M., Honneger A.M.,
 RA Hawk R., Givol D., Ulrich A., Schleisinger J.;
 RT "All autophosphorylation sites of epidermal growth factor (EGF)
 RT receptor and HER2/neu are located in their carboxyl-terminal tails.
 RT Identification of a novel site in EGF receptor."

RL J. Biol. Chem. 264:10667-10671(1989).
 RN [22]
 RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND
 RP ASN-528.
 RX MEDLINE=96398132; PubMed=8962717;
 RA Smith K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.;
 RT "Analysis of the glycosylation patterns of the extracellular domain of
 RT the epidermal growth factor receptor expressed in Chinese hamster
 RT ovary fibroblasts.";
 RL Growth Factors 13:121-132(1996).
 RN [23]
 RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND
 RP ASN-603.
 RX MEDLINE=20198209; PubMed=10731668;
 RA Sato C., Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;
 RT "Characterization of the N-oligosaccharides attached to the atypical
 RT Asn-X-Cys sequence of recombinant human epidermal growth factor
 RT receptor.";
 RL J. Biochem. 127:65-72(2000).
 Query Match 46.5%; Score 3167; DB 1; Length 1210;
 Best Local Similarity 49.8%; Pred. No. 8.3e-153; Matches 632; Conservative 176; Mismatches 346; Indels 116; Gaps 22;
 11 LIALALPPGAA--STOVCTGDMKRLRPA SPETHLMDLRHLYOGCQVQVQGNILETYLPTN 68
 14 LIALCPASRALREKVKVQGTSMKLTGCTEDHFLSLQRMFNCEVVLGNILETYVQRN 73
 69 ASLSFLQDIOEVQVYLIANQVROVPLQRIYRGVLTQFEDNVALAVLNGDPLANTTP 128
 74 YDLSFLKTIQEVAGVLIATLNTVERIPLENQIIRGNMYENSYALAVLSYND----- 126
 129 VTGASPGSLRELQRLSLLEILKGGVLIQRNPOLCYQDITLWKDIPHKXNQALTLIDTNR 188
 127 ---ANKTGLKELPMNLOEILHGA VRSSNPALCVESITQRDIYSSPLMSMSDPOHN 183
 189 SRACHPCSPMCKSGRCWSESESDCSLTRVCAGGCA-RCKGRPLTDCBQCAAGCTG 247
 184 LGSCKKDPSPCNSCWAGENQKLTIIQAQCSGRCKGKSPSDCHNCAAGCTG 243
 248 KHSDDLACLFHNSGICELHCPALVTVTDFESHPNPEGRYTRFASCVTACPYVYLTSD 307
 244 RESDCLVRKFRDEATCKDTCPLMLVNPITYQMDVNPBEGYXSPFATCVKCKCPRYVYTD 303
 308 VGSCTLVPLNQETLAEQVRCORCEKSKPCARVCYVGIMELHREVRVAVTSANIQEPAGC 367
 304 HGSCVRACGADSYEM-EEBQVRKCKCGPRKVCNGIGIBFQDSLSINATNIHFQVC 362
 368 KIFGSLAFLPESFGDPASNTAPLPQEPQLQVFETLEITGYLYISAMPDLSPLDSVFN 427
 363 TISGDDHILVVA PRGDSFTHTPPLDPOEIDILKTKVETITGFLILQAMPERRTDLHAEN 422
 428 LQVIRGILHNGAVSLTQIGISWLGRLRELGSGIALIHNHTLFCFVHTVPMQDLFR 487
 423 LEIRGRFKHQGFSLAVVSNITSIGRLKEISDGVIIISGNMLCYANTIMKKIKFG 482
 488 NPHQALLHTANRPPECCVGBGLAQQLCARHGCGPSTQCVCNSQFLRSGQCVBECVYL 547
 483 TSGQTKTILISNGEESCAYATQVCHALCSPGSCWPEBRDCCSNVSRGBECVDKML 542
 548 QGLPREYVNAHCLPCHPECOPONGSVTCFPEADQCVACAHYKDPFCVARCPSGVNPD 607
 543 EGRREFVBNBECIOCHBCLPQAMNITCTGRPNICQACHYIDGPHCVKTCRPAWGMB 602
 608 LSYMPIKFPDEBEGACQPCPINCTHSQVDDDKGCPABQBPASPLTISIAVVG---ILIV 664
 603 NNTLL-VMKYADAGVCHLCHPNCTYGCYGPGLBGCPTNGPKIP--SIATGMVGAALLLLIV 659
 665 VVLGVVFGIILIKRROQKIRKXTMRLLQETELVBLTPSGAMPNOAKRIKETEELRYK 724
 660 VALIGG---LPMRRRHVYKKTLLRLLOERELVEBLTTPSGAPNALILIKELBPKIK 716
 725 VLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPRANKELIDEAAYVMAGVGSPPVSR 784

Db	Accession	Protein Name	Length
Db	717	VLGSGAFGVKGLMIPGEKKYKLVATKELEATSPRANBEIDEAYVMSVDPHCR	776
Qy	785	LLGICLTSTVQLVLTQMLPMEYGCILLDHVRENRGRISQDILNNCMQITAKMSYLEDRLVHR	844
Db	777	LLGICLTSTVQLITQLMFGCILDVYVRHKONISQYILNNCVQJAKGMVLEBRILVHR	836
Qy	845	DLAARNVLVKSBNHYKITDPGLARLLDIDETFYHADGGKVPITKMALESITIRRRFTHSD	904
Db	837	DLAARNVLVKTPQHKYKITDPGLAKLIGABEKEVYHAGGKVPITKMALESITIRRYTHSD	896
Qy	905	VMSYGVTVWELMTFGAKPYDGI PAKEIPDLLEKGRLLPQPICTIDVYMIWVKCMIDSE	964
Db	897	VMSYGVTVWELMTFSKPYDGI PAEISISITLKGSRLLPQPICTIDVYMIWVKCMIDAD	956
Qy	965	CRPRRELVSESRNARDPQRFVVIQ-NEDLGPASPLDSTFYRSITLEDMDGDLVDAEY	1023
Db	957	SRPKRELLIERSKMARDPQRLVYIQSBERMHLPTDTSNFRALMDEEDMDVDVDAEY	1016
Qy	1024	LVPOQGFPCPDPAFGAGGVHHRHSSSTRSGGDLTLGLEPSEEARSPLABEGAGS	1083
Db	1017	LIPQGGFP-----SSPSTRTEPLSLISATS	1042
Qy	1084	DVFDGDLGMGAAGKIGSLPTHDPSLQRYSEDPYTPLESER--DGYVAFLCSQPEEYVN	1144
Db	1043	N-NSTVACIDRANGIQSCPIKEDSFLQRYSSDPFGALTEDSIDDTFL-----PVPEYIN	1094
Qy	1142	QPDVAPQPPSPREGPLPAAPAPAGATL-----ERPKTLSPKNGVYKDVFAFGAVENPEY	1166
Db	1095	Q-----SVF-KRPAQSVQNVYVNIQNPALNAPASKDPHQD--PSTVAGNEX	1138
Qy	1197	L-TPQGAAPQPHPPAPSPAFDNLVYWDQ-----DP-----PERGAPSPSTFKG	1239
Db	1139	LVTVQ-----PTCVNSTFSDSPAHAQKSHQISLSDNDYQDQDFPEAKENGIFKG	1189
Qy	1240	TPTAENPEYL 1249	
Db	1190	S-TAENAEYL 1198	
RESULT 12			
Q8MIL6_PIG PRELIMINARY; PRT: 1203 AA.			
AC	Q8MIL6		
DT	01-OCT-2002 (Tremblrel. 22, Created)		
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)		
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)		
DE	Epidermal growth factor receptor.		
OS	Sus scrofa (Pig).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euteria; Laurasiatheria; Cetartiodactyla; Suidae;		
OC	Sus.		
OX	NCBI_TaxID=96823;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Kim J.G., Vallet J.L., Nonneman D., Christenson R.K.;		
RL	Submitted (JUN-2002) to the EMBL/Genbank/DBJ databases.		
DR	EMBL: AY117054; AA077472.1; -, mRNA.		
DR	HSSP: Q9HC9; 1M17.		
DR	SMR; Q8MIL6; 25-525, 27-636.		
DR	GO: GO:0016020; C:membrane; IEA.		
DR	GO: GO:0005524; F:ATP binding; IEA.		
DR	GO: GO:0005006; F:epidermal growth factor receptor activity; IEA.		
DR	GO: GO:0004872; F:receptor activity; IEA.		
DR	GO: GO:0006466; P:protein amino acid phosphorylation; IEA.		
DR	GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.		
DR	InterPro: IPR006211; EGF_R_L.		
DR	InterPro: IPR006212; Furin-like.		
DR	InterPro: IPR000719; Prot_Kinase.		
DR	InterPro: IPR001245; Tyr_Kinase.		
DR	InterPro: IPR008266; Tyr_kinase_AS.		
DR	Pfam: PF00757; Furin-like; 1.		

	Query Match	Best Local Match	Similarity	Score	DB 2	Length	1209
		Match	632	Conservative	178	Mismatches	344
						Indels	117
						Gaps	23
Dr	Pfam, PF01030, Recep_L domain, 2						
Dr	PRK00109; TIRKINASE.						
Dr	Prodcm; PD000001; Prot_kinase, 1.						
Dr	SMART; SM00261; FU, 5.						
Dr	SMART; SM00219; Tyrc; 1.						
Dr	PROSITE; PS00107; PROTEIN KINASE ATP, 1.						
Dr	PROSITE; PS00011; PROTEIN KINASE DOM, 1.						
Dr	PROSITE; PS00109; PROTEIN_KINASE_TYR, 1.						
Kw	Receptor.						
Seq	SEQUENCE						
	1209 AA; 133531 MW; 2683BFB1E3690F CRC64;						
Qy	46.3%; Score 3152.5; DB 2; Length 1209;						
Db	49.7%; Pred. No. 4.6e-152;						
	Matches 632; Conservative 178; Mismatches 344; Indels 117; Gaps 23						
Qy	12 LIAL-----PGAASTVCTGTDMLRLPASBETHIDMLRHLYOGCQVYQGNILETYL						
Db	11 LLALIAHFQSPALPEEKVCQGGISNKLTOGCTFEDHLSIQRMFMNCEVVLGNILETYM						
Qy	66 PTNALSFLDIOIQVQGYVLIANNQVQVPLQRLRYRGTOLFEDNYLALVINDGPLN						
Db	71 QNSYNSLFLKTIQVAGVYVLIALTNTVEKIPLENQIIRGNVLYENTHALVLSN-----						
Qy	126 TTPYTGASPGELRELOLRSLTEILKGGVLIQRNQLCYQDTILMKDIPEHKNQALALTID						
Db	125 ----YGANKTGLRELPMNLQELIIGAFRFSNNPLCAESIQRMDYNSDPLSNBDF						
Qy	186 TNRSRACHPCSPMKSGSRCWSESEDCOSLTRVYCAAGCA-RCKGPLPTDCCHEQCAAG						
Db	181 QSGQSGCPKCPDGGGLNSCWAGKENGQKLTKVICAQCSGRCHGRSPDCCHQCAAGC						
Qy	245 TGRPHSDCLAHNNHSGICEHLCPALVTYNTDTFESMNPDEGRYTFGASCVTACPYNYL						
Db	241 TGPRESDCIVCRFRDEATCKDTCPPLMLYNPTTYQMDVNPILGXKSFQATCVKCCPRMYV						
Qy	305 STDVSGCLIVCLPLNQEYLTADGQROREKCKPCARVCYGLGMELREAVAVTSANIQEF						
Db	301 VTDHSGSCRRASSDSYEV-EEDGYRCKCKCGPGKCKNGSIGEFKQTLTSLNATNIKHF						
Qy	365 AGCKKIRFSLAFLPESFPDGPASNTAPLOPEQLOVFTLEETITGYLIYIAMPDLSPLDSV						
Db	360 RNCISISGDHLILVNAFRGDSFTTPTPLDPELDILKTVKSIITGELLIQANPEKRTGLHA						
Qy	425 FQNLQVIRGRILHNGAVSLTQGLGISLWGLRSIRELGSGLAIHHNTHLCFHTVPMDQ						
Db	420 FENLEIIRGRKQHGQESLAVAGDIASLGRSLKEISDGVYIVISGNNLCTYANTISWKK						
Qy	485 LFRPHQALHTARPEDECVGBELACHQICARHCWGRPGTQCNCSQFLRGQSVCEC						
Db	480 LFGTASQCTKTIINRSEKCEKAMHICNPICSSSGCWGPEPRDCSCNFRSGKEVCKC						
Qy	545 RVLQGLPREVYNAHNCILPCHEPCOPONGSVYCFEPREADQCYACAHHYKDPPECVAPCSGS						
Db	540 NVLEGEEREPENNAECVQCHPECLPQAKNYTCMGRGPDSCYRCHAYIDGPHCVKTCRPGI						
Qy	605 KPDLSYMPIKFPDEBGAQCPPICTHSCYDLDKCGCPAQRASPLTSTISAVY-GILL						
Db	600 AGENSTL-IWKFADANHYCHLCHPNCITYGCVGPLEGCAVDPRKPIP--SIATGIVGILL						
Qy	664 VVVLGVVFGILLIKRRQCKIRKTYMRRLLOEITELVEPLTPSGAMNQAMRLIKETELRKV						
Db	657 AVVVALGGLFLRRR-NHVRKTRILRLQEBELVEPLTPSGEAPNQALRLIKETEFKCV						
Qy	724 KVLGSGAFGYTKIMIPDGENYKIPVALIKYLRENTSPPKAKELIDEAYVAGVSPYVS						
Db	716 KVLGSGAFGYTKIMIPGEKVKIPVALIKYLRATSPKAKELIDEAYVNASVDNPIVC						
Qy	784 RLILGICLTSTVQVLTQMLPYGCLLDHYRENRGRGLGSDLLNWCQIAKMSYLEVDRLVH						
Db	776 RLILGICLTSTVQVLTQMLPYGCLLDHYRENRGRGLGSDLLNWCQIAKMSYLEVDRLVH						
Qy	844 RDLAARNVLYKSPNVKKTIDFGLARLLIDETETHADGGRPIKMMALESTILRRARFQOS						

FT	SIGNAL	1	24	Potential.
FT	CHAIN	25	1210	Epidermal growth factor receptor.
FT	TOPO_DOM	25	647	Extracellular (Potential).
FT	TRANSMEM	648	670	Potential
FT	TOPO_DOM	671	1210	Cytoplasmic (Potential).
FT	REPEAT	75	300	Approximate.
FT	REPEAT	390	600	Approximate.
FT	DOMAIN	714	981	Protein kinase.
FT	NP_BIND	720	728	ATP (By similarity).
FT	COMBIDAS	1028	1071	Ser-rich.
FT	ACT_SITE	839	839	By similarity.
FT	BINDING	747	747	ATP (By similarity).
FT	SITE	1018	1018	Important for interaction with PIK3CB
FT	MOD_RES	680	680	Phosphothreonine (by PKC) (By similarity).
FT	MOD_RES	1092	1092	Phosphotyrosine (by autocatalysis) (By similarity).
FT	MOD_RES	1110	1110	Phosphotyrosine (by autocatalysis) (By similarity).
FT	MOD_RES	1172	1172	Phosphotyrosine (by autocatalysis) (By similarity).
FT	MOD_RES	1197	1197	Phosphotyrosine (by autocatalysis) (By similarity).
FT	CARBOHYD	128	128	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	175	175	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	196	196	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	352	352	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	413	413	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	444	444	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	528	528	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	568	568	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	603	603	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	623	623	N-linked (GlcNAc. . .) (Potential).
FT	DISULFID	190	199	By similarity.
FT	DISULFID	194	207	By similarity.
FT	DISULFID	215	223	By similarity.
FT	DISULFID	219	231	By similarity.
FT	DISULFID	232	240	By similarity.
FT	DISULFID	236	248	By similarity.
FT	DISULFID	251	260	By similarity.
FT	DISULFID	264	291	By similarity.
FT	DISULFID	295	307	By similarity.
FT	DISULFID	311	326	By similarity.
FT	DISULFID	329	333	By similarity.
FT	DISULFID	506	515	By similarity.
FT	DISULFID	510	523	By similarity.
FT	DISULFID	526	535	By similarity.
FT	DISULFID	539	555	By similarity.
FT	DISULFID	558	571	By similarity.
FT	DISULFID	562	579	By similarity.
FT	DISULFID	582	591	By similarity.
FT	DISULFID	595	617	By similarity.
FT	DISULFID	620	628	By similarity.
FT	DISULFID	624	636	By similarity.
FT	CONFLICT	19	19	C -> S (in Ref. 2).
FT	CONFLICT	539	539	C -> W (in Ref. 5).
FT	CONFLICT	991	991	L -> F (in Ref. 4).
FT	CONFLICT	1116	1117	HP -> DR (in Ref. 6).
SO	SEQUENCE	1210 AA;	134853 MW;	690E20D46FD22D2F5 CRC64;
Query Match				
Best Local Similarity 49.7%; Score 3146; DB 1; Length 1210;				
Matches 635; Conservative 169; Mismatches 353; Indels 120; Gaps 24;				
QY	11	LLALLPFGAA--STOVCTGTDMDKRLPASPEHLDLMDRLHLYGCGVQVQNGIELTYLPTN	68	
DB	14	LLTALCAAGALAEKKVCGGTSTNRLTOLGTPEDFHLSLQRMVYNNCEVVLGNLEITYVGRN	73	
QY	69	ASLSPLODIOEVGVVLTAFHNOVROVPLQRLRYVGTOLFEDNVYALAVLDNGDPLANTTP	128	
DB	74	YDLSPKTIQEVAGVLTALANTVERIPLENQIINGNALYENTYALILSN-----	124	

QY	129	VTGASPGGLRELOURSLTEILKGVLLIQBNPOLCYODTIIMKDI----	184	
DB	125	-YGNTRGTLPRNLOEILIGAVRFSNNPILCNMDTIQWRDIVQVFMNSMDL---	180	
QY	185	DTNRSAPCHPESPMCKSRCKGSESDCOQLTRVVCAGGCA-RKGPLPTCCGHCQAG	243	
DB	181	-QSHSPSCPKDPSCPNKSGGGEENCQKTLKILCAQQCSHRCKRGSPPSCCHQCAG	239	
QY	244	CTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPNDEGRYTFGASCTYACPNY	303	
DB	240	CTGPREBDCLVQCFQFQDEATCKOTCPMLMVPNTTYQMDVNPBKYSFGATCVKCKPNY	299	
QY	304	LSTVSGCTVCPHINQEVTAEDGTQRCCKSKFCARCYGLGKHELRVAVTSANIQE	363	
DB	300	VTDHSGCVRAKCGPDYEV--EEDGIRCKCKDCGCRKVCNIGIGEIFGFTLSINATNIKH	358	
QY	364	PAGCKKIFGLSLAPLPESPDGPAANTPLQEOLOVFTLEITGYLYISAMPSPPLTS	423	
DB	359	FKYCTALSGDLHLLVPFKGDSFTPTPLDRELEILKTEITGFLILQAMPDWDLH	418	
QY	424	VFQNLQVIRGRILHNGAVSLTLOGLISWLGSLRLBELSGALIHNTNHLCPVHTVWD	483	
DB	419	AFENLEIIRGRTKHGQFSLAVGLNITSLGLRLKEISDGDVLIISGRNLCVANTIMX	478	
QY	484	QLFNPQHALLHTANRPDESCVSEGGLACHQCLCARGHCKGPEPTQCVNCSQFLRQECVEE	543	
DB	479	KLFETPMQKTKIMNNRBAEKDCAVNHVCNPLCSSEBGCWPBPCVSCQVNSRGECEK	538	
QY	544	CRVLQGLPREVYNNARHCLPCHPECOPONGSVTCGPEADOCVACAHYKDPFCVACRPSG	603	
DB	539	CNIIIEGPREVENSECTQCHPECLPQAMNITCTGRPDNCTIQAHYIDGHCVKTCAG	598	
QY	604	VKPLSYMPIMKPFDEBACQPCPINCTHSCVDLDDKCPABOASPLTISAVGILL	663	
DB	599	IMGNNTL-VWKYADANNVCHLCHANCTYGCAGGLQCEVWSPGKPISTATGIVGLL	657	
QY	664	VVTLGVVFGI-LIRROQKTRKTYMRLLQETELVEPLTPSGAMPNOAMILKTELRK	722	
DB	658	FIVV-VALLGIGLPMRRRHIVAKRTLRLQRELEVEPLTPSGEAPNOAHILKTEPEKK	716	
QY	723	VKVLGSGAFGVVYGIWIPGSENVKIPVAILVLEBNTPSKANKELIDAYYAVAGVSPYV	782	
DB	717	IKVLGSGAFGVVYGLWIPGSEKVIIPVALKELBEATSPKANKELIDBAYYAVASVDNHV	776	
QY	783	SRLLGICLTSTVQVLTOQIMPYGCLLDHYENRGRGLSGODLIMCQIAKGSYLEDAVL	842	
DB	777	CRLLGICLTSTVQVLTOQIMPYGCLLDYREHKNDINGQYLLNMCQIAKGNVLEDRVL	836	
QY	843	HRDIAARNVTVKSPNHYKITDPGLARLLDIDETRYHADGCVPIKMALESTILRRRFTHQ	902	
DB	837	HRDIAARNVTVKTPQHYKITDPGLAKLGAEEKYHAEGGVPIKMALESTILHRIYTHQ	896	
QY	903	SDVWSYGVTWELMTFGAKPYDGIIPAREIPBLEKGERLPORPCTIDVYIMVYKMWID	962	
DB	897	SDVWSYGVTWELMTFGSKPYDGIIPASDISILBGERLPORPCTIDVYIMVYKMWID	956	
QY	963	SECPREFRELVSEFSRMAKDPQRFVVIQ-NBDLPASPLDSTFYRSLDEMDGVDYAE	1021	
DB	957	ADSPKPRELEILERSKMAKDPQRYLVIQGDBRMKLPSPDTSNFPRLAMDDEMDVDVAD	1016	
QY	1022	EYLVPOGCFPCPDPAFGAGVTHRHNSSTRSGGDLTLGLBSEBEAPSPPLAPSGA	1081	
DB	1017	EYLVPOGCFPCPDPAFGAGVTHRHNSSTRSGGDLTLGLBSEBEAPSPPLAPSGA	1042	
QY	1082	GSDVFDGDLGWAAGKGLQSLPTHPSPLQVRSSEDPVLPSEST--DGVAVPLTSGPOPEY	1139	
DB	1043	TSN-----NSTVACTNKGSSCVKEDATLQRYSSPTGAVTEDNIDDAFL-----PVREY	1092	
QY	1140	VNCPDVVPQPPSPREGPLPARPAGATL-----EPKTLSPGKGVNDVAPFGAVENP	1194	
DB	1093	VNQ-----SWP-KRPAGSVQNPVYVHNPQLHPARGRLDHQN--PHSAVQNP	1136	
QY	1195	EYL-TPGGAAPQHPPPAPSPAFDNLTYWDQ-----DP-----PERGAPSTF	1237	

DB 1137 EYLTAAQ-----PCTLSGFSNPAIMIGKSHQMSLDNPDYQDPPFKETKENGIF 1187
QY 1238 KGTPTAENPEYIGLDVP 1254
DB 1188 KG-PTAENAEYLRVAP 1203

RESULT 14
Q5SV8_MOUSE PRELIMINARY; PRT; 1210 AA.
ID Q5SV8_MOUSE
AC Q5SV8; 01-FEB-2005 (Tremblrel. 29, Created)
DT 01-FEB-2005 (Tremblrel. 29, Last sequence update)
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
DE Epidermal growth factor receptor.
GN Name=Bgfr; ORFNames=RP23-295E4.1-001;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
CX NCBI_Taxid=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Graffham D.;
RL Submitted (FEB-2005) to the EMBL/Genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Matthews L.;
RL Submitted (FEB-2005) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL645532; CA124998.2; -; Genomic DNA.
DR EMBL; BX255277; CA135984.1; -; Genomic DNA.
DR EMBL; BX255277; CA124998.2; JOINED; Genomic DNA.
DR EMBL; AL645532; CA135984.1; JOINED; Genomic DNA.
DR SM; Q5SV8; 25-525, 26-636.
DR Ensemble; ENSMUSG0000020122; Mus musculus.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0005506; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0004674; F:protein ion binding; IEA.
DR GO; GO:0004872; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006118; F:receptor activity; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:citranemembrane receptor protein tyrosine kin. .; IEA.
DR InterPro; IPR001450; 4Fe4S_Fe_S_bd.
DR InterPro; IPR000494; EGFR_L_-like.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR002290; Set_Thr_Kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR PRINTS; PR00353; 4FESFSDOKIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_Kinase; 1.
DR SMART; SMO0261; FU; 4.
DR SMART; SMO0220; S_TK; 1.
DR SMART; SMO0219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Receptor.
SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46DE2D2F5 CRC64;

Query Match 46.2%; Score 3146; DB 2; Length 1210;
Best Local Similarity 49.7%; Freq. No. 9,8e-152;
Matches 635; Conservative 169; Mismatches 353; Indels 130; Gaps 24;
QY 11 LLLALLPGAA--STGVCTGTDMLKLPAPETHLDMRLHYGGQVVGQNLLETLPNT 68

DB 14 LITLCAAGALAEKKVCGTSTNRLLTGQGFEDHPLSLQRMYNCEVLLGNIETVYORN 73
QY 69 ASLSFLQDIQVQGVVLIANHQVQVPLQRLRYRGVLTFFEDNYALATLNDPLANTTP 128
DB 74 YDLSFLKTIQEVAGVLLALNTVERIPLENQIIRGNALYENTVALILSN----- 124
QY 129 VTGASPGGLRELOLSLREILKGGVLIQNPOLCYODITLMDI----FHKNNQALATILI 184
DB 125 -YGNRTKRLRELPNRNLOEILIGAVRPNPNPLCMMDITIQMRDLYQNVFNMSMDL---- 180
QY 185 DITNSRACHPCSPMCKSGRCWGSSEDCQSLTRVCAGCA-RCKGPLPTDCHEQCAAG 243
DB 181 -QSHSPCKPCDPCSPNCSGCGGGENCQKLTIKICAQCGSHRCGRSPDCHNQCAAG 239
QY 244 CTGPKHSCCLACLNHNSGICELHCPALVYNTVTFEEMPRPBERGRTGACVYACPYNY 303
DB 240 CTGPRESDCLVCQKQDEATCKDTCPPLMLYNPTTYQMDVNPBGKYSFATCVKCKPRNY 299
QY 304 LSTDVSGCTLVCPRLNQEVTADGTORCEKSKPCARCYGLGMBHLREVAVTSANIOE 363
DB 300 VTDHSGCVRAAGPYRYEV-BEDGIRKCKKDDGPRKYNCGNIGIGEFQDTLSINATNIK 358
QY 364 FAGCKKIFGSLAPLPESFSDPDASNTAPLOEOLQVFETLEITGYLIISAMPSPDLS 423
DB 359 FKYCTAISGDHLIPVAFKDGSPRTPLDPRELEILTKVAKEITGFLIIQAMPDWTDL 418
QY 424 VPQNLQVTRGRIINHGAVSLTLOGLIGISWGLRSLRELSGLALIHNTNLCFNTVWPD 483
DB 419 AFENLEIRIGRTKHGQPSLAVVGINLITSLGRSLKEISDDDVVISGRNLCYANTTIWK 478
QY 484 QLFRRPHOALHTANRPDECEVGEGLACHQLCARGHCNGPGPTQVCNCSQFARGQCYEB 543
DB 479 KLFETPNQKTIIMNRRAKDCKAVNVHVNPLCSSBGCGRPRDVCVQNSRGECYK 538
QY 544 CRVLQGLREBYVNAHCLPCHPECOPONGYATCFGEPAODCVACAHYKDPFCVACRCSG 603
DB 539 CNLIEGEPREVENSECIQCHPECLPQAMNITTCGRGPDNCIQCAHYIDGPHCVKTCBAG 598
QY 604 VKPDLSTYPIKFPDEBACQPCPINTCHGCVDDLDDKCKPABQASPLTSTISAVVIGLL 663
DB 599 IMGERNITL-VKTVADANNVCHCANCTYGCAGBQLOCEVWPSGPKPSIATGIVGLL 657
QY 664 VVVLGVVGEI-LIRROOKIRKYMTRRLQETELVEPLTPSGAMPQNAOMELIKETELRK 722
DB 658 FIVV-VALIGLFPERRRIVKRTLRLLQRELVLPISGEARNQAHLLIKETEKK 716
QY 723 VKVLGSGAFGVYKGIWIPDGENYKIPVALIVLRNENTSPKANKETLDEAYVAGVSPYV 782
DB 717 IKVLGSGAFGVYKGIWIPBGEKXKIPVALIKELRANSPKANKETLDEAYVAGVADNVH 776
QY 783 SRLIGICTSTVOLVTOIMPYGCLLDHYRBNRGRGLSGODLLNMCQIAKMSYLEDVLY 842
DB 777 CRLIGICTSTVOLITQIMPYGCLLDYRHKMDINGSYLLNMCQIAKMGVYLEDRLV 836
QY 843 HRDLAARVNLKSPNHYKITDQGLARLLDIDETETHADGGKVPKIMMALJESILRRRFTQ 902
DB 837 HRDLAARVNLKTPQHVKITDQGLAKLGAEEKVYHGGKVPKIMMALJESILHRIYHQ 896
QY 903 SDVMSYGVTVMLMTFGAKPYDGIIPAREIPDLREKGERLPOPICTIDVYIMVYKMWID 962
DB 897 SDVMSYGVTVMLMTFGSKPYDGIIPASDJSILKEGERLPPICTIDVYIMVYKMWID 956
QY 963 SECRPRFRELVSFSPSKARDPQRVYVIO-NEDLGPASGLDSTFRSLLEDMDKDVDAE 1021
DB 957 ADSBPKEFRELILBSPKARDQRYLVIOGDREMLPSPPTDSFYRALMDEEDMDVDAD 1016
QY 1022 EYLVPOGQFPCPDAPAPAGKVVHHRSSSTRSGGGLTGLBPSBEEBAPSPPLAPSGA 1081
DB 1017 EYLVPOGQF-----NSPST-----STPLLSLSIA 1042
QY 1082 GSDVFDGDLGKAAGKGLQSLPTHDPSPLORYSEDPVLPBSET--DGVVAPLTCSPQBY 1139

Db 1043 TSN-----NSTVACINRNGSCRVKEDAFLORYSSDFTGAVTEDNDIDAFL-----PVPEY 1092
Qy 1140 VNPDVVRQPPSPREGLPAARPAAGATL-----ERPKTLSPGKGVAVDVAFGAVENP 1194
Db 1093 VNO-----SVF-KRPAGSVONPVYHNOPHLPARERDLHYON--PHSAVGNP 1136
Qy 1195 EYL-TPQGGAAPQHPHPAPSPAFPDNLVYWDQ-----DP-----BERGAPPSTF 1237
Db 1137 EYLNTAQ-----PTCLSSGFNSPALMTQKSGSHQMSLDNPDYQODPFPEKTEPNCIF 1187
Qy 1238 KGTPAENPEYLGDDVP 1254
Db 1188 KG-PTAENAEYLRVAP 1203

RESULT 15
Q9EP98_MOUSE PRELIMINARY; PRT; 1210 AA.
ID Q9EP98_MOUSE PRELIMINARY; PRT; 1210 AA.
AC Q9EP98;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-FEB-2005 (Tremblrel. 29, Last annotation update)
DE Epidermal growth factor receptor isoform 1.
GN Name=Egfr;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C3H/101, and C57BL/6J;
RX MEDLINE=21100872; PubMed=1161793; DOI=10.1006/geno.2000.6341;
RA Beller J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
RA Schehl Sincclair C., Pearse R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D.,
RA Maibhe N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egfr transcripts encoding truncated receptor
RT isoforms";
RL Genomics 71:1-20(2001).
DR EMBL; AF275366; AAC28045.1; -; Genomic_DNA.
DR EMBL; AF275367; AAC24386.1; -; mRNA.
DR EMBL; AF275364; AAC28045.1; JOINED; Genomic_DNA.
DR EMBL; AF275365; AAC28045.1; JOINED; Genomic_DNA.
DR HSRP; O9H2C9; 1M17.
DR SMR; Q9EP98; 25-525, 27-636.
DR MGI; MGI:95284; Egfr.
DR GO; GO:0016323; C:basolateral plasma membrane; IDA.
DR GO; GO:0030139; C:endocytic vesicle; IDA.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005622; C:intracellular; IDA.
DR GO; GO:0016301; F:kinase activity; IDA.
DR GO; GO:0005515; F:protein binding; IPT.
DR GO; GO:004871; F:signal transducer activity; IDA.
DR GO; GO:0050730; F:regulation of peptidyl-tyrosine phosphorylation; IMP.
DR GO; GO:0007165; P:signal transduction; IDA.
DR InterPro; IPR001450; 4Fe4s ferredoxin.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin_repeat.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recp_L_domain; 2.
DR PRINTS; PR00353; 4FE4SFRDXIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; Fu; 4.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Receptor.
SQ SEQUENCE 1210 AA; 134840 MW; 62CD021C9DE32E18 CRC64;
Query Match 46.1%; Score 3143; DB 2; Length 1210;
Best Similarity 49.7%; Pred. No. 1,4e-151;
Matches 635; Conservative 168; Mismatches 354; Indels 120; Gaps 24;
Qy 11 LLLALPPGAA--STQVCTGTDMLRLPASPEYDMLRHLHYGCGQVVGNYLELYLEPTN 68
Db 14 LITLCAAGALBEKRTKCOGTSNRNLTLQGTEDHFLSLQRYNNCEVLLGLTEITYQRN 73
Qy 69 ASISFLQDIDQGVYLIANQVRQVPLQRLIRVGTQLFEDNYALAVLDNGDPLNTP 128
Db 74 YDLSFLKTIQEVAGYVLIANTVERIPLENQIIRGNALYENTYALALSN----- 124
Qy 129 VTGASPGGLREPLQRLSTELLKGVLLQRPQLCYQPTILMKDI-----FHKNQALTLI 184
Db 125 -YGTNRGTGLRELPRNRQDEILIGAVRSNNPILCNMDITQWRDIVQNVFMSMSMDL--- 180
Qy 185 DYNRSRACHPCSPKCKSGKSGESEDQSLTFTVCAAGCA-RCKGPLPTDCHEQCAG 243
Db 181 -QSHRSSCPKCDPCSPGSCMGSEENCQDKTLTKCAQCSHRKRGSRSDCCNQCAG 239
Qy 244 CTGPKHSDCLACLFHNSGICELHCPALVYNTDTPFESMPNBSGRYTFGASCVTACPVNY 303
Db 240 CTGPRSDCLVQCFQDEATCKDTCPPMLYNPTTYQMDVNPBEGKYSFGATCVKCPRNY 299
Qy 304 LSTVGSCTLVCPAHNDVYTAEDTORCEKSKCAVCVGLGEMHLREVAATVSANIQE 363
Db 300 VYVDHSGCVRACGPDYEV--EEDGIRCKCKDGCRCRVKNGIGIGEPDYLISATNATK 358
Qy 364 FAGCKRFGSLAFPLESFDDGPASNTAPLOEQQVETBELRTGYLISAMPSPDLPS 423
Db 359 FKCYTALSGDLHLTPVAFKDSFTRTPRDREBELIKTYKEITGFLLIQMPDMDLH 418
Qy 424 VPONLQVIRGRILHNGVSLTLQGLGISWGLRSRLBSGLALIHNTHLCPVHYVPWD 483
Db 419 AFENLEIRIGRTHQGFSLAVVGLANTITSLRLKEISDDVLIISGRNLCYANTIMWK 478
Qy 484 QLPENPQALLHTNRDEDECVBGLACHQLCARGHSCMPGPTCVNCSQPLRQECVEE 543
Db 479 KLFTPNQKTKIMNRKBEKCKAVNHVNCPLCSSESGCMGPBPCVQNSRRECEK 538
Qy 544 CRVLQGLPREYVNAHCLPCPECOFGSVYTCGPBADOCAVCAHYKDPFCVACRPSG 603
Db 539 CNLIEGPRFEVNSSEICQHPBCLPQAMNITCTGRPDNIOCAHYIDGHCVKTCGAG 598
Qy 604 VKPDLSTYMPWKPFDEEGACQPCPINCTHSCVDLDDKGCRAEGRASPLTISAVGILL 663
Db 599 IMGENNTL-VWKYDANNVCHLCHANCTYGACAGPLQGCCEVWPSGPRIPISTATIGVGLL 657
Qy 664 VVVLGVVFGI-LIRROOKIRKYMRLLOETELVEPLTBSGAMPNOAMILLKTELRK 722
Db 658 FTVV-VALGIGLEFRRHHIVRKRLRLQREIVEPLTBSGAPNOAMILLKTECEK 716
Qy 723 VKVLGSAFGTVYKGIWIPGENYKIPYAIKVLRENTSPKANKIILDEAYVMAGVSPYV 782
Db 717 IKVLGSAFGTVYKGLMIPGEKKIIPVALKELEARSPPANKIILDEAYVMASVDNPHV 776
Qy 783 SRLIGICLSTVQLVTLQMPYGLLDHVRENRLGSLQDILNMCMQIAKMSYLEDVRLV 842
Db 777 CRLLGICLSTVQLITQLMYGGCLLDVREHKDNGISQYLLNMVCQIAKMGNYLEDRLV 836
Qy 843 HRDLAANNVLYKSNHVKITDPGLARLLDIDETFYHADGGVPIKMMALBSILRRPETHQ 902
Db 837 HRDLAANNVLYKTPQHKITDPGLAKLGAEBKRYHAEGGVPIKMMALBSILHRTYHQ 896
Qy 903 SDVWSYGVYVWELMTFPKAPYDGIIPAREIPDLLEKSGRLPOPICTIDVYIMVYKCMNID 962
Db 897 SDVWSYGVYVWELMTFPSSKPYDGIIPASDISILEKSGRLPOPICTIDVYIMVYKCMNID 956
Qy 963 SECRPRELVSEFSRMAQDPQRFVVIQ-NBDLGASPLDSTFYRSLLEDDMDGLVDAAE 1021

```

Db      957  AD$RPF$RELIE$S$K$A$R$D$Q$R$T$V$I$Q$D$E$R$M$H$L$P$S$P$T$D$N$F$Y$R$A$L$M$D$E$D$M$E$D$V$D$A$D 1016
Qy      1022  EYLVPOG$F$C$P$D$P$A$P$G$M$V$H$H$R$N$S$S$T$R$S$G$G$D$T$T$G$L$E$P$S$E$E$A$P$R$S$P$A$P$S$E$G$A 1081
Db      1017  EYLTPOG$F$F$-----NSP$T$-----SRTPL$S$S$A 1042
Qy      1082  G$D$V$F$D$D$L$G$M$A$K$G$L$Q$S$L$P$T$H$D$S$P$L$Q$R$Y$S$E$D$P$T$V$P$L$P$S$E$T$--D$G$V$A$P$L$T$C$S$P$O$P$E$Y 1139
Db      1043  T$S$N$--N$S$T$V$A$C$I$N$R$N$G$S$C$R$V$K$E$D$A$F$L$Q$R$Y$S$D$P$T$G$A$V$E$D$N$I$D$A$F$L$-----F$V$P$E$Y 1092
Qy      1140  V$N$Q$P$D$V$R$P$P$P$P$R$E$G$P$L$P$A$R$P$A$G$A$T$L$----E$R$P$K$T$L$S$P$K$N$G$V$W$D$V$P$A$F$G$A$V$E$N$P 1194
Db      1093  V$N$Q$-----S$V$P$-K$R$P$A$G$S$V$Q$N$P$V$Y$H$N$Q$P$L$H$P$A$P$G$R$D$L$H$Y$Q$N$--P$H$S$N$A$V$G$N$P 1136
Qy      1195  E$Y$L$-T$P$Q$G$A$A$P$Q$P$H$P$P$A$P$S$P$A$F$D$N$L$Y$Y$W$Q$-----D$P$-----P$E$R$G$A$P$S$T$F 1237
Db      1137  E$Y$L$N$T$A$Q$-----P$T$C$L$S$G$F$N$S$P$A$L$W$I$Q$G$S$H$Q$M$S$L$D$N$P$D$Y$O$O$D$F$P$K$E$T$K$P$N$G$I$F 1187
Qy      1238  K$G$T$P$A$E$N$P$E$Y$L$G$L$D$V$P 1254
Db      1188  K$G$-P$T$A$E$N$A$E$Y$L$R$V$A$P$P 1203

```

Search completed: January 3, 2006, 11:16:43
 Job time : 267 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2006, 10:55:57 ; Search time 196 Seconds
(without alignments)
2813.369 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815

Sequence: 1 MEIALACRMGLLALLPFGA.....TFKGTPTAENPEYLGLDVPV 1255

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 244163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database :

A_Geneseq_21:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*
9: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	6815	100.0	1255 2	AAW01111 HER-2/neu
2	6815	100.0	1255 2	AAW92406 Human HER
3	6815	100.0	1255 3	AA184780 Amino aci
4	6815	100.0	1255 3	AA211198 Human HER
5	6815	100.0	1255 4	AA211198 Human HER
6	6815	100.0	1255 4	AA211198 Human HER
7	6815	100.0	1255 5	AA211198 Human HER
8	6815	100.0	1255 5	AA211198 Human HER
9	6815	100.0	1255 5	AA211198 Human HER
10	6815	100.0	1255 5	AA211198 Human HER
11	6815	100.0	1255 6	AA211198 Human HER
12	6815	100.0	1255 6	AA211198 Human HER
13	6815	100.0	1255 7	AA211198 Human HER
14	6815	100.0	1255 7	AA211198 Human HER
15	6815	100.0	1255 7	AA211198 Human HER
16	6815	100.0	1255 7	AA211198 Human HER
17	6815	100.0	1255 7	AA211198 Human HER
18	6815	100.0	1255 8	AA211198 Human HER
19	6815	100.0	1255 8	AA211198 Human HER
20	6815	100.0	1255 8	AA211198 Human HER
21	6815	100.0	1255 8	AA211198 Human HER
22	6815	100.0	1255 8	AA211198 Human HER
23	6815	100.0	1255 8	AA211198 Human HER
24	6815	100.0	1255 9	AA211198 Human HER

25	6815	100.0	1255 9	ADW95472
26	6815	100.0	1255 9	AE887742
27	6809	99.9	1255 9	ADW87390
28	6806	99.9	1255 3	AA928520
29	6806	99.9	1255 4	AA60167
30	6806	99.9	1255 4	AAE12130
31	6806	99.9	1255 5	AAE26349
32	6806	99.9	1255 5	AAE26349
33	6806	99.9	1255 5	AAU74545
34	6806	99.9	1255 6	ABR47447
35	6806	99.9	1255 6	ABP74708
36	6806	99.9	1255 6	AAE38390
37	6806	99.9	1255 6	ADA38143
38	6806	99.9	1255 7	ADA37255
39	6806	99.9	1255 7	ADB67621
40	6806	99.9	1255 8	ADH13187
41	6806	99.9	1255 8	ADM72831
42	6806	99.9	1255 8	ADO20009
43	6806	99.9	1255 8	ADQ29700
44	6806	99.9	1255 8	ABW81864
45	6806	99.9	1255 9	ADW28640

ALIGNMENTS

RESULT 1	AAW01111	standard; protein; 1255 AA.
ID	AAW01111	
AC	AAW01111	
XX		
DT	01-JAN-1997	(first entry)
XX		
DE	HER-2/neu protein.	
XX		
KM	HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;	
KW	breast cancer; ovary cancer; colon cancer; lung cancer; prostate cancer;	
XX	immunisation; tumour; vaccine; vector.	
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	Domain	676..1255
FT		/label= intracellular domain
FT		/note= "claimed domain, useful for immunisation"
XX		
PN	WO9630514-A1.	
XX		
PD	03-OCT-1996.	
XX		
PF	28-MAR-1996;	96WO-US001689.
XX		
PR	31-MAR-1995;	95US-00414417.
XX		
PA	(UNITV) UNITV WASHINGTON.	
XX		
PI	Cheever MA, Disis ML;	
XX		
DR	WPI: 1996-455361/45.	
XX		
DR	N-PSDB; AAT40739.	
XX		
PT	DNA encoding HER-2/neu poly:peptide(s) - used for prevention or treatment	
XX	of malignancies with which the HER-2/neu oncogene is associated.	
PS	Claim 2: Page 56-61; 71pp; English.	
XX		
CC	Human HER-2/neu protein (AAW01111), also called p185, or c-erbB2, is the	
CC	product of the HER-2/neu oncogene (see also AAT40739). The protein is	
CC	over-expressed in various cancers, including breast, ovarian, colon, lung	
CC	and prostate. The intracellular domain of the protein can be used to	
CC	immunise an animal against a malignancy with which the oncogene is	
CC	associated. The polypeptide can be produced in transformed host cells for	

CC use in immunisation. Alternatively, animal cells are transfected in vivo
 CC or ex vivo with a viral vector that directs expression of the polypeptide
 XX
 SQ Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 2; Length 1255;
 Beel Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELALCRGGLLALLPGLASTOYCTGDMKLRIPASPEITHLMDLRHLYGCGVQVQNL 60
DB 1 MELALCRGGLLALLPGLASTOYCTGDMKLRIPASPEITHLMDLRHLYGCGVQVQNL 60
QY 61 ELTYLPPTNASLFLDIOEVGVYLIANQVQVPLQRLIVRGTLFEDNATALAVLNG 120
DB 61 ELTYLPPTNASLFLDIOEVGVYLIANQVQVPLQRLIVRGTLFEDNATALAVLNG 120
QY 121 DPLNNTPTVTGASPGGLRELQLRLSLTEILKGGVLIQRNPQLCYODTIIMKDI FHKNNOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQLRLSLTEILKGGVLIQRNPQLCYODTIIMKDI FHKNNOLA 180
QY 181 LTLIDTNSRACHPCSPMKGSRGWGSESDCOSLTRVCAGGACRCKGPLPTDCCHQC 240
DB 181 LTLIDTNSRACHPCSPMKGSRGWGSESDCOSLTRVCAGGACRCKGPLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNBEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNBEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPRLHNOEVTABDGTQRCCKSPCARVCYGLGMEHLREVRATVSAN 360
DB 301 YNYLSTDVGSCTLVCPRLHNOEVTABDGTQRCCKSPCARVCYGLGMEHLREVRATVSAN 360
QY 361 IOEFAGCKKIFGSLAFIPESFDGDPASNTAPLOPQLOVFEITLITGLYISAPDLSLP 420
DB 361 IOEFAGCKKIFGSLAFIPESFDGDPASNTAPLOPQLOVFEITLITGLYISAPDLSLP 420
QY 421 DLVSFQNLQVIRGRILHNGAVSLTLOGLISWLGSRSLRELGSGLALIHNTHLCEFTV 480
DB 421 DLVSFQNLQVIRGRILHNGAVSLTLOGLISWLGSRSLRELGSGLALIHNTHLCEFTV 480
QY 481 FWDQLFRRPHQALHTANRBEDECVBEGIAHQLCARHGWPGPTQCVNCSQFLRGQEC 540
DB 481 FWDQLFRRPHQALHTANRBEDECVBEGIAHQLCARHGWPGPTQCVNCSQFLRGQEC 540
QY 541 VEECEVLQGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCVACAHYKDPFCVAKC 600
DB 541 VEECEVLQGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCVACAHYKDPFCVAKC 600
QY 601 PSGVPRDLSYMPIMKFPDEBEGACQPCINCTHSCVDLDDKGPABQORASPLTISIAYVG 660
DB 601 PSGVPRDLSYMPIMKFPDEBEGACQPCINCTHSCVDLDDKGPABQORASPLTISIAYVG 660
QY 661 ILVVVLGVVFGIILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOQMRILKETEL 720
DB 661 ILVVVLGVVFGIILIKRQOKIRKTYMRRLQETELVEPLTSGAMPNOQMRILKETEL 720
QY 721 RRVKVLGSGAGETVYKGIPIRGENVYKIPVAIKVIRENTPSPANKELIDBAVVMGVGSP 780
DB 721 RRVKVLGSGAGETVYKGIPIRGENVYKIPVAIKVIRENTPSPANKELIDBAVVMGVGSP 780
QY 781 YVSRLLGICLTSTVOLVTQLMFYGCLDHRNENRGLSGODLLNMCQIAKMSYLEDVR 840
DB 781 YVSRLLGICLTSTVOLVTQLMFYGCLDHRNENRGLSGODLLNMCQIAKMSYLEDVR 840
QY 841 LVHRDLAARNVLYKSPNHYKITDFGLARLLIDETBYTAADGKVPKIMWALESLIRRF 900
DB 841 LVHRDLAARNVLYKSPNHYKITDFGLARLLIDETBYTAADGKVPKIMWALESLIRRF 900
QY 901 HOSDWSYGVTVWELMTFGAKPYDGI PAEIPDLLEKGRSLQPPICITIDVMIWVKCM 960
DB 901 HOSDWSYGVTVWELMTFGAKPYDGI PAEIPDLLEKGRSLQPPICITIDVMIWVKCM 960

```

```

QY 961 IDSECRPRFRELVEFSRMAPDPQRFVVIQHEDIGASPPLDSTYRSLLEDDMDGLVDA 1020
DB 961 IDSECRPRFRELVEFSRMAPDPQRFVVIQHEDIGASPPLDSTYRSLLEDDMDGLVDA 1020
QY 1021 EBYLVPOQGFPCPPAPAGCMVHRRSSSTRSGGDLTGLPSEFEARSPAPASEG 1080
DB 1021 EBYLVPOQGFPCPPAPAGCMVHRRSSSTRSGGDLTGLPSEFEARSPAPASEG 1080
QY 1081 AGSDVFDLDMGAAGLQSLPTHDPSLQRYSEDPTVPLPSETDGYVAPLTCSPOPEYV 1140
DB 1081 AGSDVFDLDMGAAGLQSLPTHDPSLQRYSEDPTVPLPSETDGYVAPLTCSPOPEYV 1140
QY 1141 NQPDVPRQPSPRGSPPLPAPAPAGATLERPKTSLPGKXGVKDYAPAGGAVENBEYLTPO 1200
DB 1141 NQPDVPRQPSPRGSPPLPAPAPAGATLERPKTSLPGKXGVKDYAPAGGAVENBEYLTPO 1200
QY 1201 GGAAPQHPHPPAPSPADNLYYWDODPPERAPASTPGTGTANPEYLGIDVPV 1255
DB 1201 GGAAPQHPHPPAPSPADNLYYWDODPPERAPASTPGTGTANPEYLGIDVPV 1255

RESULT 2
AAW92406
ID AAW92406 standard; protein; 1255 AA.
XX
AC AAW92406;
XX
DT 21-APR-1999 (first entry)
XX
DE Human HER-2/neu oncogene protein.
XX
KW HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
  malignancy; treatment; tumour.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 676..1255
FT /note="region which elicits immune response"
XX
PN US5869445-A.
PD 09-FEB-1999.
XX
PF 01-APR-1996; 96US-00625101.
XX
PR 17-MAR-1993; 93US-00033644.
PR 12-AUG-1993; 93US-00106112.
PR 31-MAR-1995; 95US-00414417.
XX
PA (UNIW ) UNIV WASHINGTON.
PI Cheever MA, Disis ML;
XX
DR WPI; 1999-152835/13.
DR N-PSDB; AAX01912.
XX
PT Use of HER-2/neu polypeptides - for eliciting an immune response to an
  HER-2/neu associated malignancy, particularly for treating or preventing
  tumours.
XX
PS Claim 3; Col 31-38; 26pp; English.
XX
CC This sequence represents the human HER-2/neu oncogene protein. A fragment
  of this protein is used in a method for eliciting or enhancing an immune
  response to HER-2/neu protein. The polypeptide can stimulate T cells and
  B cells to produce an immune response to the HER-2/neu protein. The
  CC method can be used for immunisation against a malignancy in which the HER
  -2/neu oncogene is associated and in the treatment of an existing tumour,
  or to prevent tumour occurrence or reoccurrence
XX
SQ Sequence 1255 AA;

```

Query Match	100.0%; Score 6815; DB 2; Length 1255;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 1255; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MELAIACRWGLLALLPFGAASSTQVCTGDMKRLRLPASPETHLDMRLHYGCGQVQGNL	60
DB	1 MELAIACRWGLLALLPFGAASSTQVCTGDMKRLRLPASPETHLDMRLHYGCGQVQGNL	60
QY	61 ELTYLPPTANSLSFLODIOEVQGVVLIANQVQVPLQRLRIYRGQOLFEDNALAVLDNG	120
DB	61 ELTYLPPTANSLSFLODIOEVQGVVLIANQVQVPLQRLRIYRGQOLFEDNALAVLDNG	120
QY	121 DELANNTPTYGASPGGLRELQRLSLTEIKGGVLIQRNPQLCYQDTILMKDIFHKNNQLA	180
DB	121 DELANNTPTYGASPGGLRELQRLSLTEIKGGVLIQRNPQLCYQDTILMKDIFHKNNQLA	180
QY	181 LTLIDTNRSRACHPCSPMKSGSRCWGESSEDCQSLTRTVCAAGCARCKGPLPTDCHEQC	240
DB	181 LTLIDTNRSRACHPCSPMKSGSRCWGESSEDCQSLTRTVCAAGCARCKGPLPTDCHEQC	240
QY	241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDESMNPBGRYTFGASCTYACP	300
DB	241 AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDESMNPBGRYTFGASCTYACP	300
QY	301 YNYLSTDVGSCTVCPLEHNOEVTAEEDGTORCEKCSKPCARVCYGLGMEHLREYRAVTSAN	360
DB	301 YNYLSTDVGSCTVCPLEHNOEVTAEEDGTORCEKCSKPCARVCYGLGMEHLREYRAVTSAN	360
QY	361 IOEPAGCKKIFGSLAEFPSPFDGPASNTAPLOPEQLQVETLEITGYLYISAWPDSLIP	420
DB	361 IOEPAGCKKIFGSLAEFPSPFDGPASNTAPLOPEQLQVETLEITGYLYISAWPDSLIP	420
QY	421 DLSVFQNLQVIRGRILHNGAYSLTLOGLGISWLGSLARELSGLAIHNNHLCVHATV	480
DB	421 DLSVFQNLQVIRGRILHNGAYSLTLOGLGISWLGSLARELSGLAIHNNHLCVHATV	480
QY	481 PMDOLFRNPHOALLHTANRPEDECVGEGLACHQLCARGHCMWGPPTQVCNCSOFLRGQEC	540
DB	481 PMDOLFRNPHOALLHTANRPEDECVGEGLACHQLCARGHCMWGPPTQVCNCSOFLRGQEC	540
QY	541 VEECRVLOGLPREYVNAHCLPCHPECOFONGSVTCFGEADQCVACAHYKDPFVCVAC	600
DB	541 VEECRVLOGLPREYVNAHCLPCHPECOFONGSVTCFGEADQCVACAHYKDPFVCVAC	600
QY	601 PEGVPRDLSYMPWKPEDEGACQPCPINCTHSCVDLDKGCPEACRAPLTISIAYVG	660
DB	601 PEGVPRDLSYMPWKPEDEGACQPCPINCTHSCVDLDKGCPEACRAPLTISIAYVG	660
QY	661 ILLVVVLGVVFGILLIKRROOKIRKYTMRLLOETELVEPLTPSGAMPNOAKRIKETEL	720
DB	661 ILLVVVLGVVFGILLIKRROOKIRKYTMRLLOETELVEPLTPSGAMPNOAKRIKETEL	720
QY	721 RKVKVLGSGAFGVYKGIWIPDENYKIPVAIKVLENTSPKANKELIDEAAYMAGVGP	780
DB	721 RKVKVLGSGAFGVYKGIWIPDENYKIPVAIKVLENTSPKANKELIDEAAYMAGVGP	780
QY	781 YVSRLLGICLTSTVQVLTQMLPBGCLLDHYRENRGLSGODLLNMCQIAKMSYLEDVR	840
DB	781 YVSRLLGICLTSTVQVLTQMLPBGCLLDHYRENRGLSGODLLNMCQIAKMSYLEDVR	840
QY	841 LVHRDLAARNVLYKSPNHVKITDPGLARLLIDETETHADGGKVPKIMMALESILRRRT	900
DB	841 LVHRDLAARNVLYKSPNHVKITDPGLARLLIDETETHADGGKVPKIMMALESILRRRT	900
QY	901 HOSDVSXGVTWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVYIMVCKM	960
DB	901 HOSDVSXGVTWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVYIMVCKM	960
QY	961 IOSECPRRRELVSERSMARDPQRFVYIIONEDLGASPLDSTFFYSLLLEDDMGDLVDA	1020
DB	961 IOSECPRRRELVSERSMARDPQRFVYIIONEDLGASPLDSTFFYSLLLEDDMGDLVDA	1020
QY	1021 BEYLVPQGFPCPDPAFGAGVHHHRSSSTRSGGDLTLGLEPSEEBAPRSPLAPSEG	1080

DB	1021 BEYLVPQGFPCPDPAFGAGVHHHRSSSTRSGGDLTLGLEPSEEBAPRSPLAPSEG	1080
QY	1081 AGSDVFDGDLGMAKAGLQSLTPTHDSPLOQRSEDPYPLPSETGGYVAPLCSQPQEVY	1140
DB	1081 AGSDVFDGDLGMAKAGLQSLTPTHDSPLOQRSEDPYPLPSETGGYVAPLCSQPQEVY	1140
QY	1141 NOPDVRPQPPSPREGPLPAAPAGATLEBPKTSLSGKGVWDVFAFGAVENPEYLTPO	1200
DB	1141 NOPDVRPQPPSPREGPLPAAPAGATLEBPKTSLSGKGVWDVFAFGAVENPEYLTPO	1200
QY	1201 GGAAPQHPHPAPFAPDNLVYWDODPPERGAPPESTFGKTPAENPEYGLDVPV	1255
DB	1201 GGAAPQHPHPAPFAPDNLVYWDODPPERGAPPESTFGKTPAENPEYGLDVPV	1255

RESULT 3
AA584780
ID AA584780 standard; protein; 1255 AA.

XX
AC AA584780;

DT 08-AUG-2000 (first entry)

XX Amino acid sequence of the SPLICE erdb-2 receptor protein.

XX SPLICE erdb-2 receptor protein; cell transformation disorder; cancer;
XX tumor cell proliferation; tissue degeneration; arthropathy;
XX bone resorption; inflammatory disease; degenerative disorder;
XX wound healing.

OS Homo sapiens.

PN W0200020579-A1.

PD 13-APR-2000.

PF 01-OCT-1999; 99WO-CA000912.

PK 02-OCT-1998; 98US-00165192.

XX (UYMC-) UNIV MCMMASTER.

PA Muller WJ, Siegel PM;

PI WPI; 2000-303768/26.

DR N-PSDB; AAA14812.

XX Nucleic acid encoding an erdb 2 receptor protein designated SPLICE erdb-
PT 2, inhibitors of the protein are useful for treatment of cancer.

PS Claim 3; Fig 2; 60pp; English.

XX The present sequence represents a SPLICE erdb-2 receptor protein. The
CC protein has an in-frame deletion of 16 amino acids, 2 of which are
CC conserved cysteine residues, compared to the unspliced protein. The erdb-
CC 2 polynucleotide is used to construct probes for detecting disorders of
CC cell transformation such as cancer. Antibodies to the protein may be used
CC to detect SPLICE erdb-2 in a sample. Agents (e.g. antisense
CC oligonucleotides) which inhibit the expression of SPLICE erdb-2 are
CC useful for reducing tumor cell proliferation and treating cancer.
CC Substances which stimulate SPLICE erdb-2 are useful for treating
CC conditions involving damaged cells including conditions in which
CC degeneration of tissue occurs, such as arthropathy, bone resorption,
CC inflammatory diseases, degenerative disorders of the central nervous
CC system and wound healing

XX Sequence 1255 AA:

Query Match 100.0%; Score 6815; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MELAAICRNGILLALLP	PGAASTOVCTGTDMKRLR	PASPEHILDMLEHLYGCGVOYGNL	60																																																			
Db	1	MELAAICRNGILLALLP	PGAASTOVCTGTDMKRLR	PASPEHILDMLEHLYGCGVOYGNL	60																																																			
QY	61	ELUYLPTNMSLSF	LODIOEVQGYVL	IHHNOVROVPLRLR	VRGSTOLFEDUYALAVDNG	120																																																		
Db	61	ELUYLPTNMSLSF	LODIOEVQGYVL	IHHNOVROVPLRLR	VRGSTOLFEDUYALAVDNG	120																																																		
QY	121	DPLNNTTPVTGAS	PGGLRELOBSL	TEILKGVVLIQBNPOL	CYODITLMKDI	IFHKNQOLA	180																																																	
Db	121	DPLNNTTPVTGAS	PGGLRELOBSL	TEILKGVVLIQBNPOL	CYODITLMKDI	IFHKNQOLA	180																																																	
QY	181	LTIIDTNRSPACHPC	SPMCKGSR	CWSSSEDCSL	TTTVCAAGCARCKG	PLPTCCHQC	240																																																	
Db	181	LTIIDTNRSPACHPC	SPMCKGSR	CWSSSEDCSL	TTTVCAAGCARCKG	PLPTCCHQC	240																																																	
QY	241	ANGCTBPKSDCLACH	FHNHSGICE	LHCPLVTYNTPTFESMP	PEGRYTFGASC	VTAC	300																																																	
Db	241	ANGCTBPKSDCLACH	FHNHSGICE	LHCPLVTYNTPTFESMP	PEGRYTFGASC	VTAC	300																																																	
QY	301	YNYLSTDVGS	CTLVCP	LHNOEVTAE	DGTORCEKSK	ECARCYAG	MEHLREVAVTSAN	360																																																
Db	301	YNYLSTDVGS	CTLVCP	LHNOEVTAE	DGTORCEKSK	ECARCYAG	MEHLREVAVTSAN	360																																																
QY	361	IOEFACCKI	FGSLAF	PBSFDG	PASNTAPLO	BOUQVET	TEITGYLTSAMP	DSLP	420																																															
Db	361	IOEFACCKI	FGSLAF	PBSFDG	PASNTAPLO	BOUQVET	TEITGYLTSAMP	DSLP	420																																															
QY	421	DLSVPNL	LOYIRRI	LHNGAYSL	TLOGLG	SWLGRSL	RELGSGLAL	IHNTHLC	CFNHTV	480																																														
Db	421	DLSVPNL	LOYIRRI	LHNGAYSL	TLOGLG	SWLGRSL	RELGSGLAL	IHNTHLC	CFNHTV	480																																														
QY	481	PMQDLFRN	PHQALHT	TANR	PEDE	CVGBGLACHOL	CARHGCW	GPPTOC	CVNCSQPL	RGEC	540																																													
Db	481	PMQDLFRN	PHQALHT	TANR	PEDE	CVGBGLACHOL	CARHGCW	GPPTOC	CVNCSQPL	RGEC	540																																													
QY	541	VEBCRYL	OGILPRE	VYNAHCL	PC	HBCOPONGS	YTCG	PEADOC	CAAHYKD	PPFC	VIAR	600																																												
Db	541	VEBCRYL	OGILPRE	VYNAHCL	PC	HBCOPONGS	YTCG	PEADOC	CAAHYKD	PPFC	VIAR	600																																												
QY	601	PSGVKRDLS	WPIWK	PDEEGA	QCP	PCIN	CHSC	VDLDDG	CPA	EO	RASPLT	SI	SVNG	660																																										
Db	601	PSGVKRDLS	WPIWK	PDEEGA	QCP	PCIN	CHSC	VDLDDG	CPA	EO	RASPLT	SI	SVNG	660																																										
QY	661	ILLVNV	GVFGILIKR	ROOKIR	KRYTYR	RLLO	ETEL	VEPLTP	SG	AMP	NO	MR	LIK	TEL	720																																									
Db	661	ILLVNV	GVFGILIKR	ROOKIR	KRYTYR	RLLO	ETEL	VEPLTP	SG	AMP	NO	MR	LIK	TEL	720																																									
QY	721	RKYKVL	GS	GA	FG	TVYKGI	WI	PD	GENV	KI	PVAIK	Y	RENT	SP	KANKEIL	DEAY	V	MA	VG	SP	780																																			
Db	721	RKYKVL	GS	GA	FG	TVYKGI	WI	PD	GENV	KI	PVAIK	Y	RENT	SP	KANKEIL	DEAY	V	MA	VG	SP	780																																			
QY	781	YYSRL	L	G	I	C	T	S	T	V	O	L	T	O	L	M	P	Y	G	C	L	L	D	H	R	E	N	G	R	I	G	S	O	D	L	M	W	C	M	O	I	A	G	M	S	Y	L	D	V	R	840					
Db	781	YYSRL	L	G	I	C	T	S	T	V	O	L	T	O	L	M	P	Y	G	C	L	L	D	H	R	E	N	G	R	I	G	S	O	D	L	M	W	C	M	O	I	A	G	M	S	Y	L	D	V	R	840					
QY	841	L	V	H	R	D	L	A	A	R	N	V	L	K	S	P	M	H	V	K	I	T	T	P	G	L	A	R	L	D	I	D	E	T	E	H	A	D	G	K	P	I	K	M	A	L	E	S	I	L	R	R	F	T	900	
Db	841	L	V	H	R	D	L	A	A	R	N	V	L	K	S	P	M	H	V	K	I	T	T	P	G	L	A	R	L	D	I	D	E	T	E	H	A	D	G	K	P	I	K	M	A	L	E	S	I	L	R	R	F	T	900	
QY	901	H	O	S	D	V	M	S	Y	G	T	V	M	E	L	M	T	F	G	A	K	P	Y	D	I	E	P	A	R	E	I	P	D	L	L	E	K	E	R	L	P	O	P	I	C	T	I	D	V	M	I	M	V	K	M	960
Db	901	H	O	S	D	V	M	S	Y	G	T	V	M	E	L	M	T	F	G	A	K	P	Y	D	I	E	P	A	R	E	I	P	D	L	L	E	K	E	R	L	P	O	P	I	C	T	I	D	V	M	I	M	V	K	M	960
QY	961	IOSECBPR	FREL	Y	SE	SR	MA	R	D	P	O	R	F	V	I	O	N	E	D	I	G	P	A	S	P	L	D	S	T	Y	R	S	L	E	B	D	D	M	G	D	I	V	D	A	1020											
Db	961	IOSECBPR	FREL	Y	SE	SR	MA	R	D	P	O	R	F	V	I	O	N	E	D	I	G	P	A	S	P	L	D	S	T	Y	R	S	L	E	B	D	D	M	G	D	I	V	D	A	1020											
QY	1021	BEVLV	Y	Q	O	G	F	F	C	D	P	A	G	A	M	V	H	R	R	S	S	T	R	G	G	D	L	T	G	L	P	E	S	E	B	A	R	S	P	L	A	B	S	G	1080											
Db	1021	BEVLV	Y	Q	O	G	F	F	C	D	P	A	G	A	M	V	H	R	R	S	S	T	R	G	G	D	L	T	G	L	P	E	S	E	B	A	R	S	P	L	A	B	S	G	1080											
QY	1081	AGSDVE	D	G	D	I	G	M	A	K	I	O	G	L	P	T	H	D	S	P	L	O	R	Y	S	E	D	P	V	P	L	S	E	F	N	D	G	V	A	P	L	T	C	S	P	O	E	Y	1140							

[illegible]

```

Db      61  ELTVLPYMASLSFLQDIOEVQGVLLAHNQVROVPLQRLIRVRGTQLFEDNYALAVLDNG  120
Qy      121  DPLNNTTPTGASPGSLRELQLSLLEILKGVLLIORNOLCYOPTILMKDIFHKNOLA  180
Db      121  DPLNNTTPTGASPGSLRELQLSLLEILKGVLLIORNOLCYOPTILMKDIFHKNOLA  180
Qy      181  LTLITNRSRACHPCSPMKGSRCMGSESEDCSLTRTVCAAGACARCKPLPTDCHEQC  240
Db      181  LTLITNRSRACHPCSPMKGSRCMGSESEDCSLTRTVCAAGACARCKPLPTDCHEQC  240
Qy      241  AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTPESMPNDEGRYTFGASCYTACP  300
Db      241  AAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDTPESMPNDEGRYTFGASCYTACP  300
Qy      301  YNYLTSDVGSCTLVCPHNOEVTABDGTORCKSCPCARVCYGLGMEHLREVRATSN  360
Db      301  YNYLTSDVGSCTLVCPHNOEVTABDGTORCKSCPCARVCYGLGMEHLREVRATSN  360
Qy      361  IOEPAGCKKIFGSLAFLPESPDGDPASNTAPLOPELOVFEITLGYLISAMPDILP  420
Db      361  IOEPAGCKKIFGSLAFLPESPDGDPASNTAPLOPELOVFEITLGYLISAMPDILP  420
Qy      421  DLSVFQNLQVIRGRILHNGAYSLLTQGLGISWLGRLSRLRELSGLALIHNTHLCEVHTV  480
Db      421  DLSVFQNLQVIRGRILHNGAYSLLTQGLGISWLGRLSRLRELSGLALIHNTHLCEVHTV  480
Qy      481  PWDQLFRRPHQALLTANRPEDECEGEGLAGHQLCARHGMNRPPOCNCSQFLRGQGC  540
Db      481  PWDQLFRRPHQALLTANRPEDECEGEGLAGHQLCARHGMNRPPOCNCSQFLRGQGC  540
Qy      541  VEECEVLOGLPREAVYNARHCLPCHPECPONGSVTCFGEALQCVACAHAHYKDPPECVANC  600
Db      541  VEECEVLOGLPREAVYNARHCLPCHPECPONGSVTCFGEALQCVACAHAHYKDPPECVANC  600
Qy      601  PEGVPRDLSEYMPIMKPEDEBGAQCPPICTHSQVLDLDDKCPAEGRASPLTISIISAVVG  660
Db      601  PEGVPRDLSEYMPIMKPEDEBGAQCPPICTHSQVLDLDDKCPAEGRASPLTISIISAVVG  660
Qy      661  ILLVVLGVVREGILLKRRQOKIRKTYMRLLQETELVEPLTSGAMPNQAORILKETEL  720
Db      661  ILLVVLGVVREGILLKRRQOKIRKTYMRLLQETELVEPLTSGAMPNQAORILKETEL  720
Qy      721  RKVYVLSGSAFGTVYKGIWIPDGENYKIPVAIKVIRENTSPRANKELIDEAYVMAGVSP  780
Db      721  RKVYVLSGSAFGTVYKGIWIPDGENYKIPVAIKVIRENTSPRANKELIDEAYVMAGVSP  780
Qy      781  YVSRLLGICLTSTVQLVTLQMPYGLLDHVRNRRGLSQDILLNMQIAKGMSTYLEDVR  840
Db      781  YVSRLLGICLTSTVQLVTLQMPYGLLDHVRNRRGLSQDILLNMQIAKGMSTYLEDVR  840
Qy      841  LVHRDLAARNVLKSPNHYKITDFGLARLLDIDETRYHADGKVPKMMALSSILARRFT  900
Db      841  LVHRDLAARNVLKSPNHYKITDFGLARLLDIDETRYHADGKVPKMMALSSILARRFT  900
Qy      901  HOSDWWSYGVTWELMTFGAKXPYDGI PAKEIPDLLEKGERLPOPICTIDYVMIMVCKM  960
Db      901  HOSDWWSYGVTWELMTFGAKXPYDGI PAKEIPDLLEKGERLPOPICTIDYVMIMVCKM  960
Qy      961  IDSECRPRELVSERSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSLLEDDMDGLVDA  1020
Db      961  IDSECRPRELVSERSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSLLEDDMDGLVDA  1020
Qy      1021  EBYLVPOQGFCCPDPAFGAGVNHHRSSSTRSGGGLTLGLEPSEEAAPSPPLAPSG  1080
Db      1021  EBYLVPOQGFCCPDPAFGAGVNHHRSSSTRSGGGLTLGLEPSEEAAPSPPLAPSG  1080
Qy      1081  AASDVFDGLGMAAGLQSLPHTDPSPLQRYSEDPTVLPSEITDQYVAPLCSQPEYV  1140
Db      1081  AASDVFDGLGMAAGLQSLPHTDPSPLQRYSEDPTVLPSEITDQYVAPLCSQPEYV  1140
Qy      1141  NOPDVPRQPSPREGBLPAARPAAGATLERPKTILSPGKNCVNDVAFGAVENPEVLTQ  1200
Db      1141  NOPDVPRQPSPREGBLPAARPAAGATLERPKTILSPGKNCVNDVAFGAVENPEVLTQ  1200

```

```

Qy      1201  GGAAPQHPHPPAFSPAFDNLYYMDODPPERGAPPESTFKGTFRANPEYLGLDVPV  1255
Db      1201  GGAAPQHPHPPAFSPAFDNLYYMDODPPERGAPPESTFKGTFRANPEYLGLDVPV  1255

RESULT 5
AAG88267
ID  AAG88267 standard; protein; 1255 AA.
XX
AC  AAG88267;
XX
DT  11-SEP-2001 (first entry)
XX
DE  HER2/neu amino acid sequence.
XX
KW  Human; HER2/neu; epitope; human leukocyte antigen; HLA; T cell;
KW  immune response; vaccine; tumour; cancer; cytostatic; immunostimulant;
KW  tumour-associated antigen; T lymphocyte; cytotoxic T lymphocyte; CTL.
XX
OS  Homo sapiens.
XX
PN  WO200141787-A1.
XX
PD  14-JUN-2001.
XX
PF  11-DEC-2000; 2000WO-US033591.
XX
PR  10-DEC-1999; 99US-00458299.
XX
PA  (EPIIM-) EPIIMUNE INC.
PI  Fikes J, Sette A, Sidney J, Southwood S, Chesnut R, Celis E;
PI  Keogh B;
DR  WPI: 2001-374995/39.
XX
PT  An isolated prepared HER2/neu epitope useful in a vaccine for inducing
PS  cellular immune responses for the prevention and treatment of cancer.
XX
PS  Disclosure; Page 15; 1999p; English.
XX
CC  The present invention describes isolated prepared HER2/neu epitopes (I).
CC  Also described are: (1) a clonal cytotoxic T lymphocyte (CTL) that is
CC  culture in vitro and binds to a complex of an epitope (II), bound to a
CC  human leukocyte antigen (HLA) molecule; (2) a peptide (II) comprising (I)
CC  and a second epitope and the peptide is less than 50 contiguous amino
CC  acids that have 100% identity with a native peptide sequence of HER2/neu;
CC  (3) a vaccine composition (III) comprising (II) and a pharmaceutical
CC  excipient; (4) an isolated nucleic acid encoding a peptide comprising (I)
CC  and (5) an isolated nucleic acid encoding (II). (I) has cytostatic and
CC  immunostimulant activities, and can be used in vaccines. (II), (II) and
CC  (III) are useful for inducing cellular immune responses for the
CC  prevention and treatment of cancer. (I) and (II) are useful for
CC  monitoring or evaluating an immune response to a tumour-associated
CC  antigen when incubated with a T lymphocyte sample form a patient and
CC  detecting the presence of bound T lymphocyte to (I) or (II). Epitope
CC  based vaccines mean that immunosuppressive epitopes that may be present
CC  in whole antigens may be avoided. Selected epitopes may be combined to
CC  enhance immunogenicity. The possible pathological side effects caused by
CC  infectious agents or whole protein antigen is eliminated. The vaccine
CC  provides the ability to direct and focus an immune response to multiple
CC  selected antigens from the same pathogen. Epitope-based anti-tumour
CC  vaccines provides the opportunity to combine epitopes derived from
CC  multiple tumour-associated molecules addressing the problem of tumour-
CC  tumour variability and reducing the likelihood of tumour escape due to
CC  antigen loss. AAG88266 to AAG89121 represent amino acid sequences used in
CC  the exemplification of the present invention
XX
SQ  Sequence 1255 AA:

```

Query Match 100.0%; Score 6815; DB 4; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MELALCRWGLLALLPPGAAS	TOVCTGTDMLRLPAS	PERHLDMLRLHYGCCV	QVQGNL	60
Db	1	MELALCRWGLLALLPPGAAS	TOVCTGTDMLRLPAS	PERHLDMLRLHYGCCV	QVQGNL	60
Qy	61	ELTYLPTNASISFL	ODIOEVQGVYLIAHNO	VOVPLQRLIRV	GTQLFEDNYALAVL	120
Db	61	ELTYLPTNASISFL	ODIOEVQGVYLIAHNO	VOVPLQRLIRV	GTQLFEDNYALAVL	120
Qy	121	DLNNTTAVTAS	PGSLRELQRLITELIK	GGVLIQRNPQLCYQ	DTILMKDIFHNNQDLA	180
Db	121	DLNNTTAVTAS	PGSLRELQRLITELIK	GGVLIQRNPQLCYQ	DTILMKDIFHNNQDLA	180
Qy	181	LTLIDTNBSRACH	PCSPMCKSGRCWSES	SEDQSLTRFYCAGG	CARCKGPLPTDC	240
Db	181	LTLIDTNBSRACH	PCSPMCKSGRCWSES	SEDQSLTRFYCAGG	CARCKGPLPTDC	240
Qy	241	AAGCTGPRHSDCL	ACLFHNSGICELACP	ALVTYNTDTPES	MNPEGRTFGASC	300
Db	241	AAGCTGPRHSDCL	ACLFHNSGICELACP	ALVTYNTDTPES	MNPEGRTFGASC	300
Qy	301	YNYISTDVGSCT	LVCPHNOETAE	DGTQRCCKSPCAR	VCYGIGMEHLREVR	AVTSAN
Db	301	YNYISTDVGSCT	LVCPHNOETAE	DGTQRCCKSPCAR	VCYGIGMEHLREVR	AVTSAN
Qy	361	IOEFAGCKKI	FGSLAFLPES	FDGPASNTAPLQ	PQQLQVETL	EEITGILYISAMP
Db	361	IOEFAGCKKI	FGSLAFLPES	FDGPASNTAPLQ	PQQLQVETL	EEITGILYISAMP
Qy	421	DISVFQNTQVIR	GRILLHNGAVSLT	QGLISWLGIS	WLSRLS	RELGSGLALIHNT
Db	421	DISVFQNTQVIR	GRILLHNGAVSLT	QGLISWLGIS	WLSRLS	RELGSGLALIHNT
Qy	481	PMDOLFRPHQAL	LHTANRBE	CEVGEGLACHQL	CARGCMGPGPT	CCVNCQSOLF
Db	481	PMDOLFRPHQAL	LHTANRBE	CEVGEGLACHQL	CARGCMGPGPT	CCVNCQSOLF
Qy	541	VEECVAVLQGL	PREVYNARHCL	PCHECOPQNGSVT	CFGEADQCA	CAHYKDP
Db	541	VEECVAVLQGL	PREVYNARHCL	PCHECOPQNGSVT	CFGEADQCA	CAHYKDP
Qy	601	PSGVPRDIS	YMPIMKFPDE	BEGACQCPIN	CTHSCVDL	DKGCPADQ
Db	601	PSGVPRDIS	YMPIMKFPDE	BEGACQCPIN	CTHSCVDL	DKGCPADQ
Qy	661	ILLVVAVG	VFGLIKRRQOKI	RKYTMRLQ	ETELVEPLTP	SGAMPQAO
Db	661	ILLVVAVG	VFGLIKRRQOKI	RKYTMRLQ	ETELVEPLTP	SGAMPQAO
Qy	721	RKVAVLGS	GAFGTVKGIW	IPDGENVKI	PVAIKV	RENTSPKANKEI
Db	721	RKVAVLGS	GAFGTVKGIW	IPDGENVKI	PVAIKV	RENTSPKANKEI
Qy	781	YVSRLLGIC	LSTVQLMPY	GCLLDH	RENRLQSODIL	LNKCMQIAKMS
Db	781	YVSRLLGIC	LSTVQLMPY	GCLLDH	RENRLQSODIL	LNKCMQIAKMS
Qy	841	LVRHLA	RNVLYKSPNHVKI	MDPGLARLDI	DEFEYHAD	GSKVIXKMALES
Db	841	LVRHLA	RNVLYKSPNHVKI	MDPGLARLDI	DEFEYHAD	GSKVIXKMALES
Qy	901	HOSDW	SVGVTVWELMT	FGAKPYDGI	PAEIPDL	LEKGERLPO
Db	901	HOSDW	SVGVTVWELMT	FGAKPYDGI	PAEIPDL	LEKGERLPO
Qy	961	IDSECR	PRPRELVS	EFSSMAR	DPQRFVVIQ	NEDLPAS
Db	961	IDSECR	PRPRELVS	EFSSMAR	DPQRFVVIQ	NEDLPAS
Qy	1021	BEYLV	POQGF	PCPDPA	PGAGV	HRHRSS
Db	1021	BEYLV	POQGF	PCPDPA	PGAGV	HRHRSS

Qy	1081	AGSDV	FDGLGMAKGIQSL	PTHDPSP	LQYSSDP	TVLP	LPSE	TDGYA	PLTCS	QPEYV	1140
Db	1081	AGSDV	FDGLGMAKGIQSL	PTHDPSP	LQYSSDP	TVLP	LPSE	TDGYA	PLTCS	QPEYV	1140
Qy	1141	NQPDVR	POPSPSR	RGPLPA	APAGATLER	PRTLS	SPGK	NGVYKDV	FAFGA	VENPEYL	1200
Db	1141	NQPDVR	POPSPSR	RGPLPA	APAGATLER	PRTLS	SPGK	NGVYKDV	FAFGA	VENPEYL	1200
Qy	1201	GGAAP	QHPHPPA	FPAPF	ADNL	LYWPD	PERGAP	STFKGT	PTAEN	PEYL	1255
Db	1201	GGAAP	QHPHPPA	FPAPF	ADNL	LYWPD	PERGAP	STFKGT	PTAEN	PEYL	1255

Qy	1081	AGSDV	FDGLGMAKGIQSL	PTHDPSP	LQYSSDP	TVLP	LPSE	TDGYA	PLTCS	QPEYV	1140
Db	1081	AGSDV	FDGLGMAKGIQSL	PTHDPSP	LQYSSDP	TVLP	LPSE	TDGYA	PLTCS	QPEYV	1140
Qy	1141	NQPDVR	POPSPSR	RGPLPA	APAGATLER	PRTLS	SPGK	NGVYKDV	FAFGA	VENPEYL	1200
Db	1141	NQPDVR	POPSPSR	RGPLPA	APAGATLER	PRTLS	SPGK	NGVYKDV	FAFGA	VENPEYL	1200
Qy	1201	GGAAP	QHPHPPA	FPAPF	ADNL	LYWPD	PERGAP	STFKGT	PTAEN	PEYL	1255
Db	1201	GGAAP	QHPHPPA	FPAPF	ADNL	LYWPD	PERGAP	STFKGT	PTAEN	PEYL	1255
RESULT 6											
AAB85458											
ID AAB85458 standard; protein, 1255 AA.											
XX											
AC AAB85458;											
XX											
DT 25-SEP-2001 (first entry)											
XX											
DE Human HER-2/neu protein.											
XX											
KW Antigen-presenting cell; immunogenic; immune response; HER-2/neu; oncogene; cancer; cytostatic; vaccine; p185; c-erbB2.											
XX											
OS Homo sapiens.											
PN WO200153463-n2.											
XX											
PD 26-JUL-2001.											
XX											
PF 19-JAN-2001; 2001WO-US001850.											
XX											
PR 21-JAN-2000; 2000US-0177545P.											
XX											
PA (CORI-) CORIXA CORP.											
XX											
PI Cheever MA, Hand-Zimmermann S;											
XX											
DR WPI, 2001-476112/51.											
XX											
DR N-PSDB; AAH23392.											
XX											
PT New antigen-presenting cells, useful as vaccines for eliciting or enhancing an immune response to HER-2/neu protein, particularly useful for treating or preventing cancer, e.g. breast cancer.											
XX											
PS Claim 2; Page 41-46; 49pp; English.											
XX											
CC The invention provides an isolated antigen-presenting cell, which expresses at least an immunogenic portion of a polypeptide that produces an immune response to HER-2/neu protein. The antigen-presenting cells are useful as vaccines for eliciting or enhancing an immune response to HER-2/neu protein, particularly in treating or preventing malignancies in which the HER-2/neu oncogene is associated. Specifically, these are useful for treating or preventing cancer, e.g. breast cancer, ovarian, colon, lung or prostate cancers. The present sequence represents the human HER-2/neu protein (also known as p185 or c-erbB2)											
XX											
SQ Sequence 1255 AA;											
Query Match 100.0%; Score 6815; DB 4; Length 1255;											
Best Local Similarity 100.0%; Pred. No. 0;											
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;											
Qy	1	MELALCRWGLLALLPPGAAS	TOVCTGTDMLRLPAS	PERHLDMLRLHYGCCV	QVQGNL	60					
Db	1	MELALCRWGLLALLPPGAAS	TOVCTGTDMLRLPAS	PERHLDMLRLHYGCCV	QVQGNL	60					
Qy	61	ELTYLPTNASISFL	ODIOEVQGVYLIAHNO	VOVPLQRLIRV	GTQLFEDNYALAVL	120					
Db	61	ELTYLPTNASISFL	ODIOEVQGVYLIAHNO	VOVPLQRLIRV	GTQLFEDNYALAVL	120					

QY 121 DPLNNTTGVTSFGSLRELAQLRSLTEILKGGVLIQRNPOLCYOPTILMKDIFHKNNOLA 180
 DB 121 DPLNNTTGVTSFGSLRELAQLRSLTEILKGGVLIQRNPOLCYOPTILMKDIFHKNNOLA 180
 QY 181 LTLIDTNSRACHPCS PMCKGSRGWSESESDCSLTRTVACAGCARCKGRLPTDCHBQC 240
 DB 181 LTLIDTNSRACHPCS PMCKGSRGWSESESDCSLTRTVACAGCARCKGRLPTDCHBQC 240
 QY 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNBEGRYTFASCVTACP 300
 DB 241 AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNBEGRYTFASCVTACP 300
 QY 301 YNYLSTDVSGCTLVCPLNHOBTAEADGTORCEKSKPCARVYCGIGMEHLREVRVAVTSAN 360
 DB 301 YNYLSTDVSGCTLVCPLNHOBTAEADGTORCEKSKPCARVYCGIGMEHLREVRVAVTSAN 360
 QY 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPEOLQVFETLBEITGYLYISAMPDSL 420
 DB 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPEOLQVFETLBEITGYLYISAMPDSL 420
 QY 421 DLSVQNLQVTRGRILHNGAYSLTIOGIGISWLGRLSRLBELSGSLALIHNTHLCPVHTV 480
 DB 421 DLSVQNLQVTRGRILHNGAYSLTIOGIGISWLGRLSRLBELSGSLALIHNTHLCPVHTV 480
 QY 481 PMDOLFRPHOALHTANRPEDECVBEGLAHQLCARGHCMGPGPTOCVNGSQFLRGQBC 540
 DB 481 PMDOLFRPHOALHTANRPEDECVBEGLAHQLCARGHCMGPGPTOCVNGSQFLRGQBC 540
 QY 541 VBECEVLQGLPREVYNABHCLPCHPECOFQNGSVTCFGBEADQVACAHYKDPPECVAR 600
 DB 541 VBECEVLQGLPREVYNABHCLPCHPECOFQNGSVTCFGBEADQVACAHYKDPPECVAR 600
 QY 601 PSQVPRDLSYMPYKFPBEBGACQPCPINTCHSCVDLDDKGPAPORASPLTISIAYVG 660
 DB 601 PSQVPRDLSYMPYKFPBEBGACQPCPINTCHSCVDLDDKGPAPORASPLTISIAYVG 660
 QY 661 ILLVVVLVVGFIILKRRQOKIRKTYMRLLQETLVEPLPSSGAMPPOAQRILKEITEL 720
 DB 661 ILLVVVLVVGFIILKRRQOKIRKTYMRLLQETLVEPLPSSGAMPPOAQRILKEITEL 720
 QY 721 RKVKVLSGAGFTVYKGIWIPDGENVKI PVAIKVLRENTSPYANKIILDEAVYMAVGSP 780
 DB 721 RKVKVLSGAGFTVYKGIWIPDGENVKI PVAIKVLRENTSPYANKIILDEAVYMAVGSP 780
 QY 781 YVSRLLGICLTSTVOLVQLMPYGCILDHVRENRGLSGODILANCMQOIAKMSYLBVR 840
 DB 781 YVSRLLGICLTSTVOLVQLMPYGCILDHVRENRGLSGODILANCMQOIAKMSYLBVR 840
 QY 841 LVHNDLARBNTLVKSPNHVKITDPLGLARLUDIDETRYHADGKVIKMMALSTILRRFT 900
 DB 841 LVHNDLARBNTLVKSPNHVKITDPLGLARLUDIDETRYHADGKVIKMMALSTILRRFT 900
 QY 901 HQSDWMSGVTVWELMTFGAKPYDGI PAREIPDLLEKSERLPQPICTIDVYMIWVKCM 960
 DB 901 HQSDWMSGVTVWELMTFGAKPYDGI PAREIPDLLEKSERLPQPICTIDVYMIWVKCM 960
 QY 961 IDSECRPPRELIVSFSRMAARDPQRFVYIQNEDLCPASPLDSTFYRSLLEDDMDGLVDA 1020
 DB 961 IDSECRPPRELIVSFSRMAARDPQRFVYIQNEDLCPASPLDSTFYRSLLEDDMDGLVDA 1020
 QY 1021 EBYLVPOQGFPCPDPAAGAGVHHRHSSSTRSGGDLTLGLBSEBAPRSPILABEG 1080
 DB 1021 EBYLVPOQGFPCPDPAAGAGVHHRHSSSTRSGGDLTLGLBSEBAPRSPILABEG 1080
 QY 1081 AGSDVFPDDDLGMAKGLQSLPTHDPSPLOQRYSDPTVPLPSETDGYVAPLTCSPQPEV 1140
 DB 1081 AGSDVFPDDDLGMAKGLQSLPTHDPSPLOQRYSDPTVPLPSETDGYVAPLTCSPQPEV 1140
 QY 1141 NOPDVRPOPSPREBGPPLPAAPAGATLEBPKTLSFGKGVKVDVAPFGAGAVENPEYLTPO 1200
 DB 1141 NOPDVRPOPSPREBGPPLPAAPAGATLEBPKTLSFGKGVKVDVAPFGAGAVENPEYLTPO 1200

QY 1201 GGAAPQHPHPAPFSPAFDNLVYMDQPPERGAPESTFGKPTAENPEYLGLDV 1255
 DB 1201 GGAAPQHPHPAPFSPAFDNLVYMDQPPERGAPESTFGKPTAENPEYLGLDV 1255
 RESULT 7
 AAE20479
 ID AAE20479 standard; protein; 1255 AA.
 AC AAE20479;
 DT 01-JUL-2002 (first entry)
 DE Human Her-2/neu protein.
 KW Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 human leukocyte antigen; HLA; vaccine; malignancy; cytotoxic.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Region 1021..1030 /note="Naturally processed HLA-B44-restricted epitope"
 PN WO200214503-A2.
 PD 21-FEB-2002.
 PF 14-AUG-2001; 2001WO-US041733.
 PR 14-AUG-2000; 2000US-0225152P.
 PR 28-SEP-2000; 2000US-0236428P.
 PR 21-FEB-2001; 2001US-0270520P.
 PA (CORI-) CORIXA CORP.
 PI Hand-Zimmermann S, Cheever MA, Foy TM, Lodes MD, Kalos MD;
 PI McNeill PD, Veddyck TS;
 DR WPI; 2002-280758/32.
 DR N-PSDB; AAD32743.
 PS Novel isolated Her-2/Neu polypeptide composition useful for therapy,
 prevention and diagnosis of cancer, preferably breast cancer.
 CC Disclosure: Page 114-117; 129pp; English.
 CC The invention relates to an isolated Her-2/Neu polypeptide composition
 effective for eliciting an immune response. The invention is useful for
 eliciting an immune response in a patient, where the patient is human
 leukocyte antigen (HLA)-B44 positive or is affected with breast cancer.
 CC The composition is useful for the therapy and diagnosis of cancer,
 preferably breast cancer, in pharmaceutical compositions, e.g., vaccine
 and other compositions, for stimulating and/or preventing T cells specific for
 Her-2/Neu polypeptide and for inhibiting the development of cancer in a
 patient. The invention is useful for stimulating a T cell response in a
 human patient, as probe or primer for nucleic acid hybridisation, to
 selectively form duplex molecules with complementary stretches of the
 entire Her-2/Neu gene or gene fragments of interest, to isolate a full
 length gene from a suitable library, and to direct expression of a
 polypeptide in appropriate host cells. The composition is useful in
 prophylactic or therapeutic applications and for the treatment of cancer,
 preferably for the immunotherapy of breast cancer and other Her-2/Neu-
 associated malignancies. The invention is useful in gene therapy. The
 present sequence is human Her-2/neu protein
 Sequence 1255 AA:
 Query Match 100.0%; Score 6815; DB 5; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELAALCRWGILLALLPPGAASVQVCTGTDKMLRLPASPEITHDMLRHL.YOGCQVVGNTL 60
Db 1 MELAALCRWGILLALLPPGAASVQVCTGTDKMLRLPASPEITHDMLRHL.YOGCQVVGNTL 60
QY 61 ELTYLPTNASISFLDIOIEVOGYVLI.IAHNOVQVPLQRLRIYRGQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASISFLDIOIEVOGYVLI.IAHNOVQVPLQRLRIYRGQLFEDNYALAVLDNG 120
QY 121 DPLNNTTAVTGASPGGLRELOLRSLTEILKGGVLIQORNPOLCYODPTLMKDI.FHKNNOLA 180
Db 121 DPLNNTTAVTGASPGGLRELOLRSLTEILKGGVLIQORNPOLCYODPTLMKDI.FHKNNOLA 180
QY 181 LTLIDTNSRACHPCSPWCKSGRCWSSSEDCQSILTRTVACGAGCARCGPLPTDCHEQC 240
Db 181 LTLIDTNSRACHPCSPWCKSGRCWSSSEDCQSILTRTVACGAGCARCGPLPTDCHEQC 240
QY 241 AAGCTGPRHSDCLACLFHNSGICELHCPALVTYNTDFFESMPNBEGRYTGASCVTAACP 300
Db 241 AAGCTGPRHSDCLACLFHNSGICELHCPALVTYNTDFFESMPNBEGRYTGASCVTAACP 300
QY 301 YNYLSTDVSGCTVLCPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLAEVRAVTSAN 360
Db 301 YNYLSTDVSGCTVLCPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLAEVRAVTSAN 360
QY 361 IOEFAGCKKIFGSLAFLEBSFDGDPASNTAPLOPEQLQVFETLEBITGYLYISAWPDSL 420
Db 361 IOEFAGCKKIFGSLAFLEBSFDGDPASNTAPLOPEQLQVFETLEBITGYLYISAWPDSL 420
QY 421 DLSVFNQVIRGRILHNGAVSLTQIGISWLGIRSLREIGSGIALHNNHTL.CFYHTV 480
Db 421 DLSVFNQVIRGRILHNGAVSLTQIGISWLGIRSLREIGSGIALHNNHTL.CFYHTV 480
QY 481 PMDOLFRPHQALHTANRPEDECVGEGLACHOLCARGHCGPPTQCVCNCSQFLRQEC 540
Db 481 PMDOLFRPHQALHTANRPEDECVGEGLACHOLCARGHCGPPTQCVCNCSQFLRQEC 540
QY 541 VEEBCVLOGLPREYVNAHCLPCHBECOPONGSVTCSEPREADQCAAHYKDPFCVARC 600
Db 541 VEEBCVLOGLPREYVNAHCLPCHBECOPONGSVTCSEPREADQCAAHYKDPFCVARC 600
QY 601 PSQVCPDLSSYPIWKFPEBEACQCPCLNCHSCVDLDKCPAQRASPLTSIISAVG 660
Db 601 PSQVCPDLSSYPIWKFPEBEACQCPCLNCHSCVDLDKCPAQRASPLTSIISAVG 660
QY 661 ILLVVVLGVNFGIILKRQOKIRKTYMRLLQETLVEPLTPSGAMPQAQMRILKETEL 720
Db 661 ILLVVVLGVNFGIILKRQOKIRKTYMRLLQETLVEPLTPSGAMPQAQMRILKETEL 720
QY 721 RKVVLLSGGAFGTVYKGIWIPDGENVKI.PVAIKVIRENTSPKANKIIDEAYVNAVGSP 780
Db 721 RKVVLLSGGAFGTVYKGIWIPDGENVKI.PVAIKVIRENTSPKANKIIDEAYVNAVGSP 780
QY 781 YVSRLLGICLTSTQVLTQMLPYGCLLDHVRNBRRLSODLLNMCQI.AAGMSYLEVR 840
Db 781 YVSRLLGICLTSTQVLTQMLPYGCLLDHVRNBRRLSODLLNMCQI.AAGMSYLEVR 840
QY 841 LVHRDLAARNVLYKSPNHVKTDFGLARLLDDETEYADGGKYVIKMMALESIIRRF 900
Db 841 LVHRDLAARNVLYKSPNHVKTDFGLARLLDDETEYADGGKYVIKMMALESIIRRF 900
QY 901 HQSDWVSYGTVWELMTFGAKPYDGI.PAREIPDLLEKGBRLPQPICTIDVYMIWVKWM 960
Db 901 HQSDWVSYGTVWELMTFGAKPYDGI.PAREIPDLLEKGBRLPQPICTIDVYMIWVKWM 960
QY 961 IDSECRPPRELVSFSSMAADPQRFVYIQNEDLGPAPLDSTFRRSLLEDWDMDLVA 1020
Db 961 IDSECRPPRELVSFSSMAADPQRFVYIQNEDLGPAPLDSTFRRSLLEDWDMDLVA 1020
QY 1021 EBYLVPQOGFFCPDPAFGAGVWHHRHSSSTRSGGDLTLGLBPSSEBAPRSLAPSEG 1080
Db 1021 EBYLVPQOGFFCPDPAFGAGVWHHRHSSSTRSGGDLTLGLBPSSEBAPRSLAPSEG 1080
QY 1081 AGSDVFDGLGMAKGLQSLPTHDSPLQRYSEDPVPLPSETGUYAVAPLTCSPQPEV 1140

Db 1081 AGSDVFDGLGMAKGLQSLPTHDSPLQRYSEDPVPLPSETGUYAVAPLTCSPQPEV 1140
QY 1141 NOPDVRPOPSPRRGPLPAARPAATLERPRTLSPGNKVGVQDVAFAGVAPENBEYLTPQ 1200
Db 1141 NOPDVRPOPSPRRGPLPAARPAATLERPRTLSPGNKVGVQDVAFAGVAPENBEYLTPQ 1200
QY 1201 GGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAPSTFKGTPTAENBEYLGLDVVP 1255
Db 1201 GGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAPSTFKGTPTAENBEYLGLDVVP 1255
RESULT 8
AAU77114
ID AAU77114 standard; protein, 1255 AA.
XX
AC AAU77114;
XX
DT 05-JUN-2002 (first entry)
XX
DE Human Her-2/neu polypeptide.
XX
KW Human; Her-2/neu; cytostatic; haematological malignancy; CML;
KW acute myelogenous leukaemia; AML; chronic myelogenous leukaemia; CLL;
KW chronic lymphocytic leukaemia; myeloma; non-Hodgkin's lymphoma; MDS;
KW Hodgkin's lymphoma; T cell therapy.
XX
OS Homo sapiens.
XX
PN WO200213847-A2.
XX
XX 21-FEB-2002.
PD 13-AUG-2001; 2001WO-US025408.
XX
PF 14-AUG-2000; 2000US-00638280.
XX 28-SEP-2000; 2000US-00675304.
PR
XX
PA (CORI-) CORIXA CORP.
PI Gaiger A, Cheever MA, Hand-Zimmermann S;
XX
XX WPI; 2002-280741/32.
DR N-PSDB; ABK10730.
XX
XX Inhibiting hematological malignancy development by administering
PT polypeptide comprising immunogenic portion of Her-2/neu, polynucleotide
PT encoding the polypeptide, or antigen presenting cells expressing the
PT polypeptide.
XX
PS Disclosure; Page 71-74; 74pp; English.
XX
XX The invention relates to a method for inhibiting development of
CC haematological malignancy in a patient by administering a polypeptide
CC comprising an immunogenic portion of Her-2/neu or a polynucleotide
CC encoding the polypeptide. Antigen presenting cells that express the
CC protein can also be administered. The sequences are used for inhibiting
CC development of haematological malignancy such as acute myelogenous
CC leukaemia (AML), chronic myelogenous leukaemia (CML), chronic lymphocytic
CC leukaemia (CLL), MDS, myelomas, Hodgkin's lymphoma and non-Hodgkin's
CC lymphoma. This sequence represents the human Her-2/neu polypeptide
XX
SQ Sequence 1255 AA;
Query Match 100.0%; Score 6815; DB 5; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db      61  ELTYLPTNASLFLDIOEVGVYVLIANQVQVPLQRRLIRVGTQLFEDNYALAVLDNG 120
Qy      121  DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLTORNPOLCYOPTIIMKDI FHKNNOLA 180
Db      121  DPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLTORNPOLCYOPTIIMKDI FHKNNOLA 180
Qy      181  LTLITNNSRACHPCSPMKSGRCWGESSEDCOSLTRTYCAGGACARCKGPLPTDCHEQC 240
Db      181  LTLITNNSRACHPCSPMKSGRCWGESSEDCOSLTRTYCAGGACARCKGPLPTDCHEQC 240
Qy      241  AAGCTGPKRSDCLACLFHNSGICBLHCPALVYTNITRESMNPBEGRYTFGASCVTACP 300
Db      241  AAGCTGPKRSDCLACLFHNSGICBLHCPALVYTNITRESMNPBEGRYTFGASCVTACP 300
Qy      301  XNYLSTDVGSCTLVCPLNHQBVTABDGTORCEKSPCARVCYGLGMEHLREVRVTSAN 360
Db      301  XNYLSTDVGSCTLVCPLNHQBVTABDGTORCEKSPCARVCYGLGMEHLREVRVTSAN 360
Qy      361  IOEFAGCKKIFGSLAFLPESFDGDDPASNTAPLOPEQLOVEFTLEBITGYLISAMPDSL 420
Db      361  IOEFAGCKKIFGSLAFLPESFDGDDPASNTAPLOPEQLOVEFTLEBITGYLISAMPDSL 420
Qy      421  DLSVFQNLQVIRGRILHNQAVSLTLQGLGISWLGRLSLRELQSGGALIHNTHTLCFVHTV 480
Db      421  DLSVFQNLQVIRGRILHNQAVSLTLQGLGISWLGRLSLRELQSGGALIHNTHTLCFVHTV 480
Qy      481  PMDQLFRRPHQALHTANRPEDECVBEGJACHQLCARGHCMWPGPTOCNCQOFLRGQEC 540
Db      481  PMDQLFRRPHQALHTANRPEDECVBEGJACHQLCARGHCMWPGPTOCNCQOFLRGQEC 540
Qy      541  VEECVLQGLPREYVYNAHCLPCHEPCOPONGSVTCFGEBAQCYACAHYKDPPECVARC 600
Db      541  VEECVLQGLPREYVYNAHCLPCHEPCOPONGSVTCFGEBAQCYACAHYKDPPECVARC 600
Qy      601  PSGVPRDLSYMBIMKRPDEBGAQCPPICTHSCVDLDDKGPABQASPLTISIISAVVG 660
Db      601  PSGVPRDLSYMBIMKRPDEBGAQCPPICTHSCVDLDDKGPABQASPLTISIISAVVG 660
Qy      661  ILVVVLGVVFGILLKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKEITEL 720
Db      661  ILVVVLGVVFGILLKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKEITEL 720
Qy      721  RKVKYLGSGAFSTYKGIWIPDGENVKIPVAIKVLRENTSPKANKIIDEAVYMAVGVSF 780
Db      721  RKVKYLGSGAFSTYKGIWIPDGENVKIPVAIKVLRENTSPKANKIIDEAVYMAVGVSF 780
Qy      781  YVSRLLGICLTSTVOLVTQLMFYGCLLDHVRNENRGRLSGODLLNMCQIAKGMSTLEDR 840
Db      781  YVSRLLGICLTSTVOLVTQLMFYGCLLDHVRNENRGRLSGODLLNMCQIAKGMSTLEDR 840
Qy      841  LVHRDLAARNVLVKSPPNHVKTDPGLARLLIDETBYHADGSKVPIKMALESIIIRRPET 900
Db      841  LVHRDLAARNVLVKSPPNHVKTDPGLARLLIDETBYHADGSKVPIKMALESIIIRRPET 900
Qy      901  HOSDWASGVYTWELMTFGAKPYDGI PAEIPDLLEKGRLLPOPPCTIDVMIWVKCM 960
Db      901  HOSDWASGVYTWELMTFGAKPYDGI PAEIPDLLEKGRLLPOPPCTIDVMIWVKCM 960
Qy      961  IDSECRPFRELIVSEFSRMAWDPORFVVIQNEIDLSPASPLDSTFYRSLLEDDMDGLVDA 1020
Db      961  IDSECRPFRELIVSEFSRMAWDPORFVVIQNEIDLSPASPLDSTFYRSLLEDDMDGLVDA 1020
Qy      1021  EBYLVPQOGFCPPDPAFGAGVHHRHRSSTRSGGGDTLCLLPSEBEPAPSPPLAPSEG 1080
Db      1021  EBYLVPQOGFCPPDPAFGAGVHHRHRSSTRSGGGDTLCLLPSEBEPAPSPPLAPSEG 1080
Qy      1081  AGSDVFDGDLGMAKGLQSLPTHDSPLOQRYSEDPVPLPSETGYAVAPLCSQOPEV 1140
Db      1081  AGSDVFDGDLGMAKGLQSLPTHDSPLOQRYSEDPVPLPSETGYAVAPLCSQOPEV 1140
Qy      1141  NOPDVRPOPSPREGPLPAARPAAGATLERPKTSLSPKQGVVDVAFGGAIVENPEYLTPQ 1200
Db      1141  NOPDVRPOPSPREGPLPAARPAAGATLERPKTSLSPKQGVVDVAFGGAIVENPEYLTPQ 1200

Db      1141  NOPDVRPOPSPREGPLPAARPAAGATLERPKTSLSPKQGVVDVAFGGAIVENPEYLTPQ 1200
Qy      1201  GGAPQPPPPAFSPAFDNLVYMDODPPERGAPSTFGKPTPAENPEYLGDLVPV 1255
Db      1201  GGAPQPPPPAFSPAFDNLVYMDODPPERGAPSTFGKPTPAENPEYLGDLVPV 1255

RESULT 9
AAM51143
ID AAM51143 strand: protein; 1255 AA.
XX
AC AAM51143;
XX
DT 17-JUN-2002 (first entry)
XX
DE Human Her-2/neu oncogene-encoded p185 glycoprotein.
XX
KW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;
KW tyrosine kinase; receptor; c-erbB2; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1..653
FT /note="extracellular domain"
FT Domain 676..1255
FT /note="intracellular domain"
FT Domain 990..1255
FT /note="phosphorylation domain"
PN WO200212341-A2.
PD 14-FEB-2002.
PP 03-AUG-2001; 2001WO-US024283.
PR 03-AUG-2000; 2000US-00632507.
PA (CORI-) CORIYA CORP.
PA (SMK) SMTHTLINE BECHAM BIOLOGICALS.
PI Cheever MA, Gheysen D;
XX
DR MPI: 2002-241743/29.
DR N-PSDB; ABA92250.
XX
PT Her-2/neu fusion protein for treating or preventing cancer by eliciting
PT or enhancing an immune response to the protein, has Her-2/neu
PT extracellular domain fused to Her-2/neu intracellular or phosphorylation
PT domain.
XX
PS Claim 68; Fig 7; 141pp; English.
XX
CC The present sequence is that of human Her-2/neu (p185 glycoprotein or c-
CC erbB2), an oncogenic self-protein and target for anti-cancer vaccines.
CC The Her-2/neu gene is amplified and p185 is overexpressed in a variety of
CC cancers, including breast, ovarian, colon, lung and prostate cancer. Her-
CC 2/neu is a member of the tyrosine kinase family of receptor-like
CC glycoproteins. It comprises an extracellular domain with homology to the
CC epidermal growth factor receptor (EGFR), a highly hydrophobic
CC transmembrane domain and a C-terminal intracellular domain that also
CC shows homology to EGFR. Its overexpression correlates with a poor
CC prognosis in breast and ovarian cancers. The invention provides Her-2/neu
CC fusion proteins, nucleic acids encoding them, viral vectors, and vaccines
CC comprising the fusion proteins or nucleic acid molecules. In preferred
CC fusion proteins, the extracellular domain of a Her-2/neu protein is fused
CC to a Her-2/neu intracellular domain or phosphorylation domain (or its
CC Delaap fragment). An immune response to Her-2/neu protein is elicited or
CC enhanced by administering the fusion protein in the form of a vaccine, or
CC by transfecting cells of an animal ex vivo with a nucleic acid encoding
CC the fusion protein, and delivering the transfect cells to the animal.
CC The fusion proteins, nucleic acids, and isolated specific T-cells are
CC useful for inhibiting the development of a cancer, especially breast,

```

ovarian, colon, lung or prostate cancer in a patient. T cells that specifically react with a Her-2/neu fusion protein can be used to remove tumour cells from a sample in order to inhibit the development of cancer in a patient

Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 5; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKLRIPASPEHLDMLRHLYGCGQVQGNL 60
DB 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKLRIPASPEHLDMLRHLYGCGQVQGNL 60
QY 61 ELTYLPTNASLSFLDIDIOEVGVYLAHNOVROVLOGLRTYRGOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLDIDIOEVGVYLAHNOVROVLOGLRTYRGOLFEDNYALAVLDNG 120
QY 61 ELTYLPTNASLSFLDIDIOEVGVYLAHNOVROVLOGLRTYRGOLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLDIDIOEVGVYLAHNOVROVLOGLRTYRGOLFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILMKDI FHNQOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILMKDI FHNQOLA 180
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILMKDI FHNQOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILMKDI FHNQOLA 180
QY 181 LTLIDTNSRACHPCSPCKSGRCSGESSEDCQSLTRTVCAAGCARCGPLPTDCCHQC 240
DB 181 LTLIDTNSRACHPCSPCKSGRCSGESSEDCQSLTRTVCAAGCARCGPLPTDCCHQC 240
QY 241 AAGCGPRGSDCLALHFNHSGICELHCPALVTYNTDFESPNRSGRTYTGASCVTACP 300
DB 241 AAGCGPRGSDCLALHFNHSGICELHCPALVTYNTDFESPNRSGRTYTGASCVTACP 300
QY 301 YNYLSTDVSGCTIVCPHNOEVTAEDETQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
DB 301 YNYLSTDVSGCTIVCPHNOEVTAEDETQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
QY 301 YNYLSTDVSGCTIVCPHNOEVTAEDETQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
DB 301 YNYLSTDVSGCTIVCPHNOEVTAEDETQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
QY 361 IOEPAGCKKIRGSLAFLESPFGDPASNTAPLOPQOLOVFTLEITIGLYISAMPDLR 420
DB 361 IOEPAGCKKIRGSLAFLESPFGDPASNTAPLOPQOLOVFTLEITIGLYISAMPDLR 420
QY 361 IOEPAGCKKIRGSLAFLESPFGDPASNTAPLOPQOLOVFTLEITIGLYISAMPDLR 420
DB 361 IOEPAGCKKIRGSLAFLESPFGDPASNTAPLOPQOLOVFTLEITIGLYISAMPDLR 420
QY 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISMLGRLSRLGSGALAHNHTHLCFVHTV 480
DB 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISMLGRLSRLGSGALAHNHTHLCFVHTV 480
QY 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISMLGRLSRLGSGALAHNHTHLCFVHTV 480
DB 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISMLGRLSRLGSGALAHNHTHLCFVHTV 480
QY 481 PMDQLFRPHQALHTANRPEDECEVGEGLACHQLCARHCGWPGPTQCVNSQFLRGQEC 540
DB 481 PMDQLFRPHQALHTANRPEDECEVGEGLACHQLCARHCGWPGPTQCVNSQFLRGQEC 540
QY 481 PMDQLFRPHQALHTANRPEDECEVGEGLACHQLCARHCGWPGPTQCVNSQFLRGQEC 540
DB 481 PMDQLFRPHQALHTANRPEDECEVGEGLACHQLCARHCGWPGPTQCVNSQFLRGQEC 540
QY 541 VEECEVLOGLPREVYNAHCLPCHEPCOPONGSVTCFGEADQCVACHYKDPPECVARC 600
DB 541 VEECEVLOGLPREVYNAHCLPCHEPCOPONGSVTCFGEADQCVACHYKDPPECVARC 600
QY 541 VEECEVLOGLPREVYNAHCLPCHEPCOPONGSVTCFGEADQCVACHYKDPPECVARC 600
DB 541 VEECEVLOGLPREVYNAHCLPCHEPCOPONGSVTCFGEADQCVACHYKDPPECVARC 600
QY 601 PSGVPEDLSTYPIWKEPDEBEGACOPCPINCHSCVDLDDKGPAPORASPLTISIAVVG 660
DB 601 PSGVPEDLSTYPIWKEPDEBEGACOPCPINCHSCVDLDDKGPAPORASPLTISIAVVG 660
QY 601 PSGVPEDLSTYPIWKEPDEBEGACOPCPINCHSCVDLDDKGPAPORASPLTISIAVVG 660
DB 601 PSGVPEDLSTYPIWKEPDEBEGACOPCPINCHSCVDLDDKGPAPORASPLTISIAVVG 660
QY 661 ILVVVGLVVFGLIKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
DB 661 ILVVVGLVVFGLIKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
QY 661 ILVVVGLVVFGLIKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
DB 661 ILVVVGLVVFGLIKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
QY 721 RKVKVLGSGAGFTYKGIWIPGENVVKIPVAIKVIRENTSPKANKEIIDEAIVMAVGVSF 780
DB 721 RKVKVLGSGAGFTYKGIWIPGENVVKIPVAIKVIRENTSPKANKEIIDEAIVMAVGVSF 780
QY 721 RKVKVLGSGAGFTYKGIWIPGENVVKIPVAIKVIRENTSPKANKEIIDEAIVMAVGVSF 780
DB 721 RKVKVLGSGAGFTYKGIWIPGENVVKIPVAIKVIRENTSPKANKEIIDEAIVMAVGVSF 780
QY 781 YVSRLLGICLSTVQLVQMLPMYGCLLDHVRNRRGLSODLLNMCQIAGMSYLEBVR 840
DB 781 YVSRLLGICLSTVQLVQMLPMYGCLLDHVRNRRGLSODLLNMCQIAGMSYLEBVR 840
QY 781 YVSRLLGICLSTVQLVQMLPMYGCLLDHVRNRRGLSODLLNMCQIAGMSYLEBVR 840
DB 781 YVSRLLGICLSTVQLVQMLPMYGCLLDHVRNRRGLSODLLNMCQIAGMSYLEBVR 840
QY 841 LVHRDLAARNVLYKSPNHVKTIDFGLARLLDIDETEHADGGKVIKKMALESILRRFT 900
DB 841 LVHRDLAARNVLYKSPNHVKTIDFGLARLLDIDETEHADGGKVIKKMALESILRRFT 900
QY 841 LVHRDLAARNVLYKSPNHVKTIDFGLARLLDIDETEHADGGKVIKKMALESILRRFT 900
DB 841 LVHRDLAARNVLYKSPNHVKTIDFGLARLLDIDETEHADGGKVIKKMALESILRRFT 900
QY 901 HOSDWSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVYMLWKCM 960
DB 901 HOSDWSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVYMLWKCM 960

```

```

DB 901 HOSDWSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVYMLWKCM 960
QY 961 IDSECRPRELVESEFARMARDPQRFVYI QNEDIGPASPLDSTFYRSLLEDMDGLVDA 1020
DB 961 IDSECRPRELVESEFARMARDPQRFVYI QNEDIGPASPLDSTFYRSLLEDMDGLVDA 1020
QY 1021 EBYLVPOQGFCDPAPAGAGVHHRRSSSTRSGGDLTYGLBSEBARSPLABSEG 1080
DB 1021 EBYLVPOQGFCDPAPAGAGVHHRRSSSTRSGGDLTYGLBSEBARSPLABSEG 1080
QY 1021 EBYLVPOQGFCDPAPAGAGVHHRRSSSTRSGGDLTYGLBSEBARSPLABSEG 1080
DB 1021 EBYLVPOQGFCDPAPAGAGVHHRRSSSTRSGGDLTYGLBSEBARSPLABSEG 1080
QY 1081 AGSDVFGDDIGMAKGLQSLPTHDPSPLOKYSDDPTVPLSENDGYVAPLTCSPQPEYV 1140
DB 1081 AGSDVFGDDIGMAKGLQSLPTHDPSPLOKYSDDPTVPLSENDGYVAPLTCSPQPEYV 1140
QY 1141 AGSDVFGDDIGMAKGLQSLPTHDPSPLOKYSDDPTVPLSENDGYVAPLTCSPQPEYV 1140
DB 1141 AGSDVFGDDIGMAKGLQSLPTHDPSPLOKYSDDPTVPLSENDGYVAPLTCSPQPEYV 1140
QY 1141 AGSDVFGDDIGMAKGLQSLPTHDPSPLOKYSDDPTVPLSENDGYVAPLTCSPQPEYV 1140
DB 1141 AGSDVFGDDIGMAKGLQSLPTHDPSPLOKYSDDPTVPLSENDGYVAPLTCSPQPEYV 1140
QY 1201 GGAAPQHPHPAPFSPADNLYYWDQDPPERCAPSTFKGPTTANPEYLGLDVPY 1255
DB 1201 GGAAPQHPHPAPFSPADNLYYWDQDPPERCAPSTFKGPTTANPEYLGLDVPY 1255

```

```

RESULT 10
AAE24067
ID AAE24067 standard; protein, 1255 AA.
XX
AC AAE24067;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human Her-2 protein.
XX
KW Human; Her-2; epidermal growth factor receptor 2; infection; cancer;
KW hyperproliferative disorder; prophylaxis; inflammation; antisense;
KW tumour; gene therapy; phosphothioate backbone.
XX
OS Homo sapiens.
XX
PN MO200222636-A1.
XX
PD 21-MAR-2002.
XX
PF 12-SEP-2001; 2001WO-US028572.
XX
PR 15-SEP-2000; 2000US-00663834.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Bennett CF, Cowseart LM;
XX
DR MPI; 2002-471192/50.
XX
DR N-PSDB; AAD38904.
XX
PT Novel antisense oligonucleotide which modulates the expression of Human
PT Epidermal Growth Factor receptor, Her2, is useful for treating tumors
PT inflammation or to prevent infection in humans.
XX
PS Example 13; Page 95-107; 116pp; English.
XX
CC The invention relates to antisense compounds targeted to a nucleic acid
CC molecule encoding Her2 (human Epidermal Growth Factor receptor 2) that
CC specifically hybridizes with and inhibits the expression of Her2.
CC Antisense compounds of the invention are used for treating diseases or
CC conditions associated with Her2 such as hyperproliferative disorders e.g.
CC lung, breast, gastric, oesophageal, colon, bladder, salivary, neural or
CC cardiac cancer. They are also useful prophylactically e.g. to prevent or
CC delay infection, inflammation and tumour formation. The invention is also
CC used in gene therapy. The present sequence is human Her-2 protein
XX
SQ Sequence 1255 AA;

```

Query Match 100.0%; Score 6815; DB 5; Length 1255;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 METALACRGLLALLLPRAASTVCTGDMKLRLPASPEHLDMRLHYOCQVVGNTL 60
DB 1 METALACRGLLALLLPRAASTVCTGDMKLRLPASPEHLDMRLHYOCQVVGNTL 60
QY 61 ELTYLPTNASLFLDIOIEVGVVLIANQVQVPLQRLIRVGTOLPEDNTALAVLNG 120
DB 61 ELTYLPTNASLFLDIOIEVGVVLIANQVQVPLQRLIRVGTOLPEDNTALAVLNG 120
QY 121 DELNNTTPTVNGASPGGLRELQLRLSILTEILKGGVLIQRNQLCYQDTILMKDI FHKNNQLA 180
DB 121 DELNNTTPTVNGASPGGLRELQLRLSILTEILKGGVLIQRNQLCYQDTILMKDI FHKNNQLA 180
QY 181 LTLIDTNSRACHPCSPMKSGKSCGSESEDCOSLTRTCAGGACARCKGPLEPTDCHEQC 240
DB 181 LTLIDTNSRACHPCSPMKSGKSCGSESEDCOSLTRTCAGGACARCKGPLEPTDCHEQC 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPNPEGRTYFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMPNPEGRTYFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPPLNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRATSN 360
DB 301 YNYLSTDVGSCTLVCPPLNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRATSN 360
QY 361 IOEPFAGCKKIFGSLAFLPESFPGDPAASNTAPLOPELOVFEFLBETGLYLSAMPDILP 420
DB 361 IOEPFAGCKKIFGSLAFLPESFPGDPAASNTAPLOPELOVFEFLBETGLYLSAMPDILP 420
QY 421 DLSVQNTQVIRGRILHNGAVSLTQGLGISWLGRLSRLSELSSGLALIHNTHLFCVHTV 480
DB 421 DLSVQNTQVIRGRILHNGAVSLTQGLGISWLGRLSRLSELSSGLALIHNTHLFCVHTV 480
QY 481 FWDQLFRRPHQALHTANRPEDECVGEGLAHQLCARHCWPGPTQCVCNCSQPLRGQEC 540
DB 481 FWDQLFRRPHQALHTANRPEDECVGEGLAHQLCARHCWPGPTQCVCNCSQPLRGQEC 540
QY 541 VEECEVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCYACAHYKDPFCVAC 600
DB 541 VEECEVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCYACAHYKDPFCVAC 600
QY 601 PSGVPRDLSYMPIMKFPDEBEGACOPCINCHSCVDLDDKGPAPORASPLTSIIISAVYG 660
DB 601 PSGVPRDLSYMPIMKFPDEBEGACOPCINCHSCVDLDDKGPAPORASPLTSIIISAVYG 660
QY 661 ILLVVLGVVFGILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNQAKMRILKETEL 720
DB 661 ILLVVLGVVFGILIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNQAKMRILKETEL 720
QY 721 RKVKVLGSGAGFTVYKGIWIPGENVAKIPVAIKVIRENTSPKANKEIIDEAIVMAVGSP 780
DB 721 RKVKVLGSGAGFTVYKGIWIPGENVAKIPVAIKVIRENTSPKANKEIIDEAIVMAVGSP 780
QY 781 YVSRLLGICLTSTVOLMPYGLLDHVRNRRGLSODLLNMCQIAKMSLTEDVR 840
DB 781 YVSRLLGICLTSTVOLMPYGLLDHVRNRRGLSODLLNMCQIAKMSLTEDVR 840
QY 841 LVHRDLAARNVLYKSPNHVKTIDFGIARLLIDETEHYADGGKVDIKWMALESILRRRT 900
DB 841 LVHRDLAARNVLYKSPNHVKTIDFGIARLLIDETEHYADGGKVDIKWMALESILRRRT 900
QY 901 HOSDWSVGVYTWELMTFGAKPYDGI PAKEIIDLEKERRLPQPICTIDVYMIWVKCM 960
DB 901 HOSDWSVGVYTWELMTFGAKPYDGI PAKEIIDLEKERRLPQPICTIDVYMIWVKCM 960
QY 961 IDSECRPPRELVSFESRMARDPORFVYQNEIDLPAASPLDSTFYSRLLEDMDGLVDA 1020
DB 961 IDSECRPPRELVSFESRMARDPORFVYQNEIDLPAASPLDSTFYSRLLEDMDGLVDA 1020
QY 1021 BEYLVPQOQFPCPDPAFGAGVNHHRHSSSTRSGGDLTLGLBESBEEAPRSLAPSEG 1080
DB 1021 BEYLVPQOQFPCPDPAFGAGVNHHRHSSSTRSGGDLTLGLBESBEEAPRSLAPSEG 1080

```

```

DB 1021 BEYLVPQOQFPCPDPAFGAGVNHHRHSSSTRSGGDLTLGLBESBEEAPRSLAPSEG 1080
QY 1081 AGSDVFDGLGWAAGKGLSLPTHDPSPLQRYSEDPVPLPSETGUYAFLTCSPOPEYV 1140
DB 1081 AGSDVFDGLGWAAGKGLSLPTHDPSPLQRYSEDPVPLPSETGUYAFLTCSPOPEYV 1140
QY 1141 NOPDVRPOPSPREBGLPAARPAATLERPKTLSPGKGVVDVAFGAVENPEYLTPQ 1200
DB 1141 NOPDVRPOPSPREBGLPAARPAATLERPKTLSPGKGVVDVAFGAVENPEYLTPQ 1200
QY 1201 GGAAPQHPPPAFSPAFDNLTYWDDPPBRGAPSPSTFYGTPTAENPEYLGLDVPV 1255
DB 1201 GGAAPQHPPPAFSPAFDNLTYWDDPPBRGAPSPSTFYGTPTAENPEYLGLDVPV 1255

RESULT 11
ABR43687
ID ABR43687 standard; protein; 1255 AA.
XX
AC ABR43687;
XX
DT 29-JUL-2003 (first entry)
XX
DE Human c-erb-B-2 (ErbB-2) protein SEQ ID NO:4.
XX
KW Human; hepatocellular carcinoma associated protein; HCAP; ErbB-2;
KW c-erb-B-2; cytostatic; nephrotropic; ophthalmological; antiproliferative;
KW antirheumatic; antiarthritic; antiinflammatory; HCAP-associated disease;
KW ErbB-2 associated disease; cell hyperproliferative disease; cancer;
KW cell proliferative disease; benign proliferative disease; psoriasis;
KW glomerulonephritis; neurofibromatosis; glaucoma; rheumatoid arthritis;
KW inflammatory bowel disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..21
FT /label= signal
FT 22..1255
FT /label= C-erb-B-2
FT /note= "also known as ErbB-2"
XX
PN WO2003025228-A1.
XX
PD 27-MAR-2003.
XX
PF 18-SEP-2002; 2002WO-US029577.
XX
PR 18-SEP-2001; 2001US-0323210P.
XX
PR 09-NOV-2001; 2001US-0323250P.
XX
PA (PROT-) PROTEOLOGICS INC.
XX
PI Reiss Y, Alroy I;
XX
DR MPI; 2003-363157/34.
XX
DR N-PSDB; ACC69999.
XX
PT Treating a hepatocellular carcinoma associated protein (HCAP)-associated
PT disease, e.g. cancer or benign proliferative disease, comprises
PT administering an agent that inhibits the interaction between HCAP and
PT HCAP binding partner.
XX
PS Claim 37; Fig 5C; 59pp; English.
XX
CC The present invention describes a method for treating a hepatocellular
CC carcinoma associated protein (HCAP)-associated disease in a subject which
CC comprises administering to the subject an agent that inhibits the
CC interaction between HCAP and an HCAP binding partner. Also described is a
CC method for treating an ErbB-2 associated disease in a subject which
CC comprises administering to the subject an agent that decreases the level
CC and/or activity of HCAP in cells of a subject. HCAP has cytostatic,
CC nephrotropic, ophthalmological, antiproliferative, antirheumatic,

```

CC antiarthritic and antiinflammatory activities. The methods can be used
 CC for treating a HCAP-associated disease or an ErbB-2 associated disease,
 CC e.g. a cell (hyper)proliferative disease such as cancer, or a benign
 CC proliferative disease such as glomerulonephritis, neurofibromatosis,
 CC gliucoma, psoriasis, rheumatoid arthritis or inflammatory bowel disease.
 CC The methods can be used for treating a proliferative disorder, e.g.
 CC breast cancer, ErbB-2-related breast cancer or taxol-resistant breast
 CC cancer. The present sequence represents human c-erb-B-2 (also known as
 CC ErbB-2) from the present invention

XX Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 6; Length 1255;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELAALCWGILLALLPFGAASVQVCTGTDKLRIPASPEETHLDMRLYGGCQVQGNL 60
DB 1 MELAALCWGILLALLPFGAASVQVCTGTDKLRIPASPEETHLDMRLYGGCQVQGNL 60
QY 61 ELTYLPTNASISFLQDIOEVQGYVLIANNOYRQVPLQSLRIYRGTOLEFEDNYALAVLDNG 120
DB 61 ELTYLPTNASISFLQDIOEVQGYVLIANNOYRQVPLQSLRIYRGTOLEFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSLTELKGGVLIQRNPOLCYQDPTILMKDIFHKNNQLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSLTELKGGVLIQRNPOLCYQDPTILMKDIFHKNNQLA 180
QY 181 LTLIDTNSRACHPSCPMCKSGRCWSESSEDCQSLTRIVCAGGACRCGRLPTDCCHQC 240
DB 181 LTLIDTNSRACHPSCPMCKSGRCWSESSEDCQSLTRIVCAGGACRCGRLPTDCCHQC 240
QY 241 AAGCTGPRGSDCLACLFHNSGICELHCPALVYTYTDFEFSMPNDEGRATFGASCVTACP 300
DB 241 AAGCTGPRGSDCLACLFHNSGICELHCPALVYTYTDFEFSMPNDEGRATFGASCVTACP 300
QY 301 YNYISTDVGSCTLVCPILNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVAVTISAN 360
DB 301 YNYISTDVGSCTLVCPILNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVAVTISAN 360
QY 361 IOEPFAGCKKIFGSLAFLESFDPDASNTAPLQEPQLOVFELEETITGYLYISAMPDLP 420
DB 361 IOEPFAGCKKIFGSLAFLESFDPDASNTAPLQEPQLOVFELEETITGYLYISAMPDLP 420
QY 421 DLSVQONTQVIRGRILHNGAYSILTQGGISWLGRLSRELGSGLALIHNTHLCFVHTV 480
DB 421 DLSVQONTQVIRGRILHNGAYSILTQGGISWLGRLSRELGSGLALIHNTHLCFVHTV 480
QY 481 PWDQLFRNPQALHTANRPEDECVGEGLAHQLCARGHCWGPPTQCVCNCSQFLRGQC 540
DB 481 PWDQLFRNPQALHTANRPEDECVGEGLAHQLCARGHCWGPPTQCVCNCSQFLRGQC 540
QY 541 VEECVNLQGLREYVNAHCLPCHPECQPONGSVTCFEPHADQCAAHYDPPFCVARC 600
DB 541 VEECVNLQGLREYVNAHCLPCHPECQPONGSVTCFEPHADQCAAHYDPPFCVARC 600
QY 601 PSGVVPDLSYMPYKPEPDEGACQPCPINCTHSCVDLDDKGPAPORASPLTISAVYG 660
DB 601 PSGVVPDLSYMPYKPEPDEGACQPCPINCTHSCVDLDDKGPAPORASPLTISAVYG 660
QY 661 ILVVVLGVVFGLILKRQOQIKRKYTMRLLQETELVEPLTPSGAMPNQAQMRILKETEL 720
DB 661 ILVVVLGVVFGLILKRQOQIKRKYTMRLLQETELVEPLTPSGAMPNQAQMRILKETEL 720
QY 721 RKVVLGSGAGFTYKGIWIPGGEVVKIPVAIKVRENTSPANKEIIDEAIVMAGVGSP 780
DB 721 RKVVLGSGAGFTYKGIWIPGGEVVKIPVAIKVRENTSPANKEIIDEAIVMAGVGSP 780
QY 781 YVSRLLGICITSTVQLVQMLPYGCLDHPVRENRRGLSODLLNMCQIJAQKMSYLEYVR 840
DB 781 YVSRLLGICITSTVQLVQMLPYGCLDHPVRENRRGLSODLLNMCQIJAQKMSYLEYVR 840
QY 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLIDETEHADGKVPDKWMALESILRRFT 900

```

```

DB 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLIDETEHADGKVPDKWMALESILRRFT 900
QY 901 HQSDVMSGYVTWMLMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMIWVCKM 960
DB 901 HQSDVMSGYVTWMLMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMIWVCKM 960
QY 961 IDSECRPRFRELVEFSRMAADPQRFVYIQNEDIGPASPLDSTFYRSLLEDDMDGLVDA 1020
DB 961 IDSECRPRFRELVEFSRMAADPQRFVYIQNEDIGPASPLDSTFYRSLLEDDMDGLVDA 1020
QY 1021 EEYLVPOQGFPCPPAAGAGGMVHRRSSSTRGCGDLTLGLEPSEEARSLAPBSG 1080
DB 1021 EEYLVPOQGFPCPPAAGAGGMVHRRSSSTRGCGDLTLGLEPSEEARSLAPBSG 1080
QY 1081 AGSDVFDGDLGMAKGLQSLPTDPSPLQRYSDPTVPLPSFIDGYVAPLTCSPOPEYV 1140
DB 1081 AGSDVFDGDLGMAKGLQSLPTDPSPLQRYSDPTVPLPSFIDGYVAPLTCSPOPEYV 1140
QY 1141 NQPDVPRQPPSPRBSGPLPAARPAAGATLERPKTSLPGKNGVVKDVFARFGAVENPEYLTPQ 1200
DB 1141 NQPDVPRQPPSPRBSGPLPAARPAAGATLERPKTSLPGKNGVVKDVFARFGAVENPEYLTPQ 1200
QY 1201 GGAAPQHPHPPEAPADNLYYWDQDPPBEGAPSTKGTPTAANPEYLGLDVPV 1255
DB 1201 GGAAPQHPHPPEAPADNLYYWDQDPPBEGAPSTKGTPTAANPEYLGLDVPV 1255

```

RESULT 12
 ABR82066
 ID ABR82066 standard; protein: 1255 AA.
 AC ABR82066;
 XX
 DT 23-SEP-2003 (first entry)
 XX
 DE Human Her2/neu amino acid sequence SEQ ID NO:1.
 XX
 KW Human; Her2/neu; Her2/neu target antigen; immune response; cytostatic;
 XX immunostimulant; vaccine; cancer.
 OS Homo sapiens.
 XX
 PN WO2003055439-A2.
 XX
 PD 10-JUL-2003.
 XX
 PF 18-JUL-2002; 2002WO-US022975.
 XX
 PR 18-JUL-2001; 2001US-0306250P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Nelson EL;
 XX
 DR WPI; 2003-569400/53.
 XX
 PT New Her2/neu target antigens and polynucleotides encoding them, useful
 PT for stimulating immunoeffector cells or an immune response in a subject
 PT against cancer cells expressing Her2/neu, and for vaccination strategies.
 XX
 PS Claim 29; Fig 1A-B; 83p; English.
 XX
 CC The present invention describes a recombinant polynucleotide (1)
 CC comprising a first nucleotide sequence encoding a Her2/neu target antigen
 CC consisting of an amino acid sequence corresponding to: (a) amino acid
 CC residues 634-683 or 606-683 operatively linked to amino acid residues
 CC 1035-1255 of human Her2/neu (SEQ ID NO:1, ABR82066); or (b) amino acid
 CC residues 635-685 or 608-685 operatively linked to amino acid residues
 CC 1037-1257 of rat Her2/neu (SEQ ID NO:2, ABR82067). (1) has cytostatic and
 CC immunostimulant activities, and can be used in vaccines. The
 CC polynucleotide encoding a Her2/neu target antigen can be used for
 CC stimulating immunoeffector cells or an immune response in a subject

CC against cancer cells expressing Her2/neu. The encoded Her2/neu target
 CC antigen polypeptide is useful as a tag to detect or isolate the fusion
 CC protein containing the Her2/neu target an immune response stimulated by
 CC the Her2/neu target antigen. The Her2/neu target antigen polynucleotides
 CC and polypeptides are useful in genetic or polypeptide vaccination
 CC strategies. The present sequence represents the human Her2/neu amino acid
 CC sequence, which is used in the exemplification of the present invention

XX
 XX
 60 Sequence 1255 AA;

Query Match 100.0%; Score 6815; DB 7; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MELALCRKGLLALLPQGAASVCTGDMKRLPASPETHLDMRLHYOCQVVGSL 60
DB 1 MELALCRKGLLALLPQGAASVCTGDMKRLPASPETHLDMRLHYOCQVVGSL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGVLLAHNQVRQVPLQRLIRVGTQLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIOEVQGVLLAHNQVRQVPLQRLIRVGTQLFEDNYALAVLDNG 120
QY 121 DELNNTTPTVYGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDI FHKNNQIA 180
DB 121 DELNNTTPTVYGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDI FHKNNQIA 180
QY 181 LTLITNRRACHPGSPMKSGRCWGESSEDCQSLTRTYCAGGACRCKPLPTDCHECQ 240
DB 181 LTLITNRRACHPGSPMKSGRCWGESSEDCQSLTRTYCAGGACRCKPLPTDCHECQ 240
QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMNPBGRATFGASCVTACP 300
DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPESMNPBGRATFGASCVTACP 300
QY 301 YNYLSTDVGSCTLYVCPILHNQEVTAEDGTQRCCKSPCARVCYGLAMEHLREVAVTSAN 360
DB 301 YNYLSTDVGSCTLYVCPILHNQEVTAEDGTQRCCKSPCARVCYGLAMEHLREVAVTSAN 360
QY 361 IOEPGAGCKKIFGSLAFPPSPGDPASNTAPLOPBLQVFEFLERTGLYISAMPDSL 420
DB 361 IOEPGAGCKKIFGSLAFPPSPGDPASNTAPLOPBLQVFEFLERTGLYISAMPDSL 420
QY 421 DLSVFQNLQVIRGRILHNGAVSLTQGLGISWLGRLSRLSELSSGLALIHNNHLFCVHTV 480
DB 421 DLSVFQNLQVIRGRILHNGAVSLTQGLGISWLGRLSRLSELSSGLALIHNNHLFCVHTV 480
QY 481 PWDQLFRNHQALLHTANRPEDECVGEGLAGHQLCARHGWCPGPTQCVNCSQPLRGQEC 540
DB 481 PWDQLFRNHQALLHTANRPEDECVGEGLAGHQLCARHGWCPGPTQCVNCSQPLRGQEC 540
QY 541 VEECVLQGLPREYVNAHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVANC 600
DB 541 VEECVLQGLPREYVNAHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVANC 600
QY 601 PSGVPRDLSYMPIMKFPDEBGAQCPPICTHSQVDLDKGPABORASPLTISIISAVVG 660
DB 601 PSGVPRDLSYMPIMKFPDEBGAQCPPICTHSQVDLDKGPABORASPLTISIISAVVG 660
QY 661 ILVVVLGVVFGILIKRQOKIRKTYMRLLQETELVEBPLSPSGAMPQONRIKETEL 720
DB 661 ILVVVLGVVFGILIKRQOKIRKTYMRLLQETELVEBPLSPSGAMPQONRIKETEL 720
QY 721 RRYKVLGSGAGFTVYKGIIPDGENVKIIVALKVRENTSPYANKIIDEAVYMGVSGP 780
DB 721 RRYKVLGSGAGFTVYKGIIPDGENVKIIVALKVRENTSPYANKIIDEAVYMGVSGP 780
QY 781 YSRLLGICLTSTVOLVTOLMPYGLLDHVRNRRGLSGODLLNMCQIAKMSYLEDVR 840
DB 781 YSRLLGICLTSTVOLVTOLMPYGLLDHVRNRRGLSGODLLNMCQIAKMSYLEDVR 840
QY 841 LVHRDLAARNVAVKSPNHVKITDQGLARLLIDETRYHADGKVPKIMMALESILRRPT 900
DB 841 LVHRDLAARNVAVKSPNHVKITDQGLARLLIDETRYHADGKVPKIMMALESILRRPT 900

```

```

QY 901 HOSDWSGYVWELMTFGAKPYDGI PARFIPDLLEKGERLPQPICTIDVYIMVCKM 960
DB 901 HOSDWSGYVWELMTFGAKPYDGI PARFIPDLLEKGERLPQPICTIDVYIMVCKM 960
QY 961 IDSECRPRRELVSFSSMARDPQRFVVIQNEDLGPASPLDSTFYRSLIEDDMGDLVDA 1020
DB 961 IDSECRPRRELVSFSSMARDPQRFVVIQNEDLGPASPLDSTFYRSLIEDDMGDLVDA 1020
QY 1021 EBYLVPOGFCPPDPAAGGVHHRHSSSTRSGGDLTGLFESBEERPSPLASRG 1080
DB 1021 EBYLVPOGFCPPDPAAGGVHHRHSSSTRSGGDLTGLFESBEERPSPLASRG 1080
QY 1081 AGSDVFDGDLGAKAGLQSLPTPHDSPLOQRYSEDPVPLPSETDGYAFLTCSQPEVY 1140
DB 1081 AGSDVFDGDLGAKAGLQSLPTPHDSPLOQRYSEDPVPLPSETDGYAFLTCSQPEVY 1140
QY 1141 NQPDVRPQPPSPREGPLPAAPAGATLERPKTLSFGKGVVDVAFGAVENPEYLTPQ 1200
DB 1141 NQPDVRPQPPSPREGPLPAAPAGATLERPKTLSFGKGVVDVAFGAVENPEYLTPQ 1200
QY 1201 GGAAPQHPHPPAFSAFDNLVYWDODPBERGAPSETFGTPTAENPEYLGIDVPV 1255
DB 1201 GGAAPQHPHPPAFSAFDNLVYWDODPBERGAPSETFGTPTAENPEYLGIDVPV 1255

```

RESULT 13
 ADC09593
 ID ADC09593 standard; protein; 1255 AA.
 XX
 XX
 AC ADC09593;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Her2/Neu #SEQ ID 594.
 XX
 KW Epitope; immunological; vaccine;
 KW major histocompatibility complex class I; MHC class I; cancer;
 KW immunisation.
 OS Unidentified.
 OS
 PN WO2003008537-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 29-MAR-2002; 2002WO-US010189.
 XX
 PR 06-APR-2001; 2001US-0282211P.
 PR 07-NOV-2001; 2001US-0337017P.
 PR 07-MAR-2002; 2002US-0363210P.
 XX
 PA (CTL1-) CTL IMMUNOTHERAPIES CORP.
 XX
 PI Simard JTL, Diamond DC, Liu L, Xie Z;
 XX
 DR WPI: 2003-248010/24.
 XX
 PS Claim 1; SEQ ID NO 594; 239pp; English.
 XX
 CC The invention relates to an isolated epitope polypeptide that has high
 CC affinity for major histocompatibility complex (MHC) class I, and an
 CC epitope cluster comprising the polypeptide. Also disclosed is a vaccine
 CC or immunotherapeutic composition containing an epitope of the invention.
 CC Compositions of the invention may be used in the treatment of cancer. The
 CC method can be combined with a radiation therapy, chemotherapy,
 CC biochemotherapy or surgery. The composition is also useful for evaluating
 CC immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC
 CC-peptide complexes of the invention are useful for determining specific T

cell frequency. This method is useful for evaluating immunological response, by performing the method prior to and subsequent to an immunisation step. Compositions of the invention are useful for diagnosing a disease. The current sequence represents an epitope of the invention with high affinity for MHC class I.

Sequence 1255 AA:

Query Match 100.0%; Score 6815; DB 7; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MELALCRWGLLALLPRGAASVQCTGDMKRLPASPEHLDMLRLYQSCQVQGNL 60
DB 1 MELALCRWGLLALLPRGAASVQCTGDMKRLPASPEHLDMLRLYQSCQVQGNL 60
OY 61 ELTYLPTNASISFLODIOEVQGYVLIANQVRQVPLQRLIVRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASISFLODIOEVQGYVLIANQVRQVPLQRLIVRGTLFEDNYALAVLDNG 120
OY 121 DPLNNTPTVTGASPGGLRELQRLTEILKGGVLIQRNPOLCYODTIIMKDIFFKNQOLA 180
DB 121 DPLNNTPTVTGASPGGLRELQRLTEILKGGVLIQRNPOLCYODTIIMKDIFFKNQOLA 180
OY 181 LTLIDTNRSRACHPCSPCKSGRCWSESEDCQSLTRTVCAAGCARGKAPLPETDCHEQC 240
DB 181 LTLIDTNRSRACHPCSPCKSGRCWSESEDCQSLTRTVCAAGCARGKAPLPETDCHEQC 240
OY 241 AAGCTGPRHSDCLACLHFNHSGICELACPALVTYNTDTPESMPNREGYTGGASCVTACP 300
DB 241 AAGCTGPRHSDCLACLHFNHSGICELACPALVTYNTDTPESMPNREGYTGGASCVTACP 300
OY 301 YNYLSTDVGSCITVCPPLHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
DB 301 YNYLSTDVGSCITVCPPLHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
OY 361 IOEPAGCKKIFGSLAFLESPFDGASNTAPLQBPOLQVFETLEBITGYLIYISAMPDLSL 420
DB 361 IOEPAGCKKIFGSLAFLESPFDGASNTAPLQBPOLQVFETLEBITGYLIYISAMPDLSL 420
OY 421 DLSVQNTQVTRGRILHNGAYSILTQGISWLGIRSLRELGSALIHNTHLCPVNTV 480
DB 421 DLSVQNTQVTRGRILHNGAYSILTQGISWLGIRSLRELGSALIHNTHLCPVNTV 480
OY 481 PMDOLFRPHOALHTANRPEDECVGEGLACHQLCARHGMCPGTQCVNCSQPLRGQEC 540
DB 481 PMDOLFRPHOALHTANRPEDECVGEGLACHQLCARHGMCPGTQCVNCSQPLRGQEC 540
OY 541 VEECEVLQGLPREVYNARHCLPCHPECOFONGSVTCFGBADQCVACAHYKDPFCVARC 600
DB 541 VEECEVLQGLPREVYNARHCLPCHPECOFONGSVTCFGBADQCVACAHYKDPFCVARC 600
OY 601 PSGVPRDLSYMPIMKFPDBEGACQCPINCHSCVDLDDKGPACQASPLTSTISAVYG 660
DB 601 PSGVPRDLSYMPIMKFPDBEGACQCPINCHSCVDLDDKGPACQASPLTSTISAVYG 660
OY 661 ILLVVLGVVFGIILKRRQOKIRKTYMRILQETELVPLPSGAMPPOAQMRILKETEL 720
DB 661 ILLVVLGVVFGIILKRRQOKIRKTYMRILQETELVPLPSGAMPPOAQMRILKETEL 720
OY 721 RKVKVLGSGAFGTYKGIWIPDGENVKIPVALKVLRENTSPKANEIIDEAVYMAVGSP 780
DB 721 RKVKVLGSGAFGTYKGIWIPDGENVKIPVALKVLRENTSPKANEIIDEAVYMAVGSP 780
OY 781 YVSRLLGLCTSTVOLVQIMPYGCLLDHVRNRRRLSQDILLNCKQIAKMSYLEBVR 840
DB 781 YVSRLLGLCTSTVOLVQIMPYGCLLDHVRNRRRLSQDILLNCKQIAKMSYLEBVR 840
OY 841 LVHRDLAARNVAVKSPNHVKTIDFGIARLLDIDETEHADGGKVPKIMMALESIIARRFT 900
DB 841 LVHRDLAARNVAVKSPNHVKTIDFGIARLLDIDETEHADGGKVPKIMMALESIIARRFT 900
OY 901 HQSDVWSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVYMIWVCKWM 960

```

```

DB 901 HQSDVWSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVYMIWVCKWM 960
OY 961 IDSECRPRFRELVEFSRMRADPQRFVYIQWEDLPASPLDSTYRSLLEDMDGDLVDA 1020
DB 961 IDSECRPRFRELVEFSRMRADPQRFVYIQWEDLPASPLDSTYRSLLEDMDGDLVDA 1020
OY 1021 EBYLVPOGFPCPDPAAGAGVHRRSSSTRSGGDLTLGLPSEEBARSPAPSEB 1080
DB 1021 EBYLVPOGFPCPDPAAGAGVHRRSSSTRSGGDLTLGLPSEEBARSPAPSEB 1080
OY 1081 AGSDVFPDGLMGAKLQSLPTHDPSPLQRYSDDPVPLPSERDGYAPLTCSPQPEYV 1140
DB 1081 AGSDVFPDGLMGAKLQSLPTHDPSPLQRYSDDPVPLPSERDGYAPLTCSPQPEYV 1140
OY 1141 NOPVRRPOPSPRRGPIPAARPAQATLERPTLSPGKNQVQDVAFCGAVENPEYLTPO 1200
DB 1141 NOPVRRPOPSPRRGPIPAARPAQATLERPTLSPGKNQVQDVAFCGAVENPEYLTPO 1200
OY 1201 GGAAPQHPPPAFSPAFDNLVYMDQDPPERGAPSTFKGPTAENPEYLGDLVPV 1255
DB 1201 GGAAPQHPPPAFSPAFDNLVYMDQDPPERGAPSTFKGPTAENPEYLGDLVPV 1255

```

```

RESULT 14
ADD25484
ID ADD25484 standard; protein, 1255 AA.
XX
AC ADD25484;
XX
DT 15-JAN-2004 (first entry)
XX
DE Binding domain-immunoglobulin fusion protein-associated protein #19.
XX
KW Binding domain; immunoglobulin; fusion protein; cytostatic;
KW antidiabetic; immunosuppressive; antidiabetic; antihypertoid;
KW neuroprotective; hinge region; immunoglobulin heavy chain;
KW CH2 constant region; CH3 constant region; IgG1;
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
KW type 1 diabetes mellitus; multiple sclerosis; autoimmune disease.
XX
OS Unidentified.
XX
PN US2003118592-A1.
XX
PD 26-JUN-2003.
XX
PE 25-JUN-2002; 2002US-00207655.
XX
PF 17-JAN-2001; 2001US-0367358P.
XX
PR 17-JAN-2002; 2002US-00053530.
XX
PP 03-JUN-2002; 2002US-0385691P.
XX
PA (GENE-) GENE-CRAFT INC.
XX
PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
XX
DR WPI; 2003-801317/75.
XX
PT New binding domain-immunoglobulin fusion protein, useful for treating a
PT subject having or suspected of having a malignant condition or a B-cell
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX
PS Disclousre; SEQ ID NO 45; 157bp; English.
XX
CC The invention relates to a binding domain-immunoglobulin fusion protein
CC comprising a binding domain polypeptide that is fused to an
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
CC CH2 constant region polypeptide that is fused to the hinge region
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region
CC polypeptide that is fused to the CH2 constant region polypeptide. The

```

CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin
 CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
 CC region polypeptide, derived from (a) having 3 or more cysteine residues;
 CC where the mutated human IgG1 immunoglobulin hinge region polypeptide;
 CC contains 2 cysteine residues, where the first cysteine is not mutated; a
 CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
 CC (a) having 3 or more cysteine residues, where the mutated human IgG1
 CC immunoglobulin hinge region polypeptide contains no more than one
 CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
 CC polypeptide, derived from (a) having 3 or more cysteine residues; where
 CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains
 CC no cysteine residues. The binding domain-immunoglobulin fusion protein is
 CC capable of at least one immunological activity comprising antibody
 CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The
 CC binding domain polypeptide is capable of specifically binding to an
 CC antigen. Also included are an isolated polynucleotide encoding the
 CC binding domain-immunoglobulin fusion protein, a recombinant expression
 CC construct comprising the polynucleotide (operably linked to a promoter),
 CC a host cell transformed or transfected with a recombinant expression
 CC construct, producing the binding domain-immunoglobulin fusion protein, a
 CC pharmaceutical composition comprising the binding domain-immunoglobulin
 CC fusion protein or polynucleotide and a carrier, and treating a subject
 CC having or suspected of having a malignant condition or a B-cell disorder.
 CC The binding domain-immunoglobulin fusion protein is useful for treating a
 CC subject having or suspected of having a malignant condition or a B-cell
 CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
 CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple
 CC sclerosis or autoimmune disease. The present sequence is a binding domain
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The
 CC sequence data for this patent formed part of the printed specification
 CC and is also available in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=20030118592. The authors have not
 CC identified the sequences in the printed specification by their SEQ ID
 CC number therefore none of the sequences can be explicitly identified.

XX Sequence 1255 AA:

Query Match 100.0%; Score 6815; DB 7; Length 1255;

Best Local Similarity 100.0%; Pred. No. 0; Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELALACWGLLALLPFGAASVCTGTDMKRLPASPETHLDMRLHYOGCOVVOGNTL 60
 DB 1 MELALACWGLLALLPFGAASVCTGTDMKRLPASPETHLDMRLHYOGCOVVOGNTL 60
 QY 61 ELTYLPFNASLSFLODIOEVGVYLIANQVQVPLQRLIRYRGQLFEDNYALAVLDNG 120
 DB 61 ELTYLPFNASLSFLODIOEVGVYLIANQVQVPLQRLIRYRGQLFEDNYALAVLDNG 120
 QY 121 DPLNNTTPTVGTASPGSLRELOLRSLTEILKGGVLIQRNPOLCYOPTLTKDIFHNQOLA 180
 DB 121 DPLNNTTPTVGTASPGSLRELOLRSLTEILKGGVLIQRNPOLCYOPTLTKDIFHNQOLA 180
 QY 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCOSTLRTVCAGACARCKGPLPTDCHEQC 240
 DB 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCOSTLRTVCAGACARCKGPLPTDCHEQC 240
 QY 241 AAGCTGPRKSDCLACIHFHNSGICELHCPALVTYNTDFESMPNREGRTFGASCVTACP 300
 DB 241 AAGCTGPRKSDCLACIHFHNSGICELHCPALVTYNTDFESMPNREGRTFGASCVTACP 300
 QY 301 YNYLSTDVGSCTLVNGLNHOEYTAEDGTORCEKSKPCARVYCYGMEHLRVRVAVTSAN 360
 DB 301 YNYLSTDVGSCTLVNGLNHOEYTAEDGTORCEKSKPCARVYCYGMEHLRVRVAVTSAN 360
 QY 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLQPEOLQVFETLEBITGYLISAPDSLP 420
 DB 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLQPEOLQVFETLEBITGYLISAPDSLP 420
 QY 421 DLSVFOQLQVIRGRILHNGAYSLTLQGLISWLGRLSELGSGALLIHNTILCFVHTV 480
 DB 421 DLSVFOQLQVIRGRILHNGAYSLTLQGLISWLGRLSELGSGALLIHNTILCFVHTV 480

QY 481 PWDOLFNPQOALLHTANRPEDECVBGLACHQLCARGCWGPPTQCVCNCSQFLRGEC 540
 DB 481 PWDOLFNPQOALLHTANRPEDECVBGLACHQLCARGCWGPPTQCVCNCSQFLRGEC 540
 QY 541 VEECVVLQGLPREYVNAHNCPLCHPECOPQNSVTCFEPBADQCYACAHYKDPFCVARC 600
 DB 541 VEECVVLQGLPREYVNAHNCPLCHPECOPQNSVTCFEPBADQCYACAHYKDPFCVARC 600
 QY 601 PSGVPRDLSYMPIMKPEPBEAGCOPCINCTHSCVDLDDKGPABORASPLTSISAVVG 660
 DB 601 PSGVPRDLSYMPIMKPEPBEAGCOPCINCTHSCVDLDDKGPABORASPLTSISAVVG 660
 QY 661 ILVVVLGVFGLIKRROQIKRYTMRLLQETELVPLTPSGAMPNQAQMRILKETEL 720
 DB 661 ILVVVLGVFGLIKRROQIKRYTMRLLQETELVPLTPSGAMPNQAQMRILKETEL 720
 QY 721 RKVKVLGSGAGSTYKGIWIPGENVKIPVAIKULRENTSPRANKELIDEAYVMAVGSP 780
 DB 721 RKVKVLGSGAGSTYKGIWIPGENVKIPVAIKULRENTSPRANKELIDEAYVMAVGSP 780
 QY 781 YVSRLLGLCLSTVOLTMQLMPYGLDHYENRGRGLSQDILLNMCQIAKMSYLEBVR 840
 DB 781 YVSRLLGLCLSTVOLTMQLMPYGLDHYENRGRGLSQDILLNMCQIAKMSYLEBVR 840
 QY 841 LVHRDLAARNVLVKSNNHVKITDFGLARLDDIDETRYADGGKVDIKMMALLESILRRPT 900
 DB 841 LVHRDLAARNVLVKSNNHVKITDFGLARLDDIDETRYADGGKVDIKMMALLESILRRPT 900
 QY 901 HOSDWSYGVTVWEIMLTGAKPYDGIPIABEIPDLLEKGRLPQPICTIDVYMIWVKCM 960
 DB 901 HOSDWSYGVTVWEIMLTGAKPYDGIPIABEIPDLLEKGRLPQPICTIDVYMIWVKCM 960
 QY 961 IDSECRPFRELVSFESMARDPORFVYIQNEEDLPASPLDSTFYSLLBEDDMGDLVDA 1020
 DB 961 IDSECRPFRELVSFESMARDPORFVYIQNEEDLPASPLDSTFYSLLBEDDMGDLVDA 1020
 QY 1021 BEYLVPQGFPCPDPAAGGVVHRHRSJSTRSGGDLTLGLLEPSEBAPRSLAPSEB 1080
 DB 1021 BEYLVPQGFPCPDPAAGGVVHRHRSJSTRSGGDLTLGLLEPSEBAPRSLAPSEB 1080
 QY 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLORYSEDPYVPLPSETDGYVAPLTCSPQPEV 1140
 DB 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLORYSEDPYVPLPSETDGYVAPLTCSPQPEV 1140
 QY 1141 NOPDVPRQPSRBSGPLAARPAAGATLERPKTSLSGKGVVMDVPAFGAVENPEYLTPQ 1200
 DB 1141 NOPDVPRQPSRBSGPLAARPAAGATLERPKTSLSGKGVVMDVPAFGAVENPEYLTPQ 1200
 QY 1201 GGAAPQPHPPAFSPAFDNLVYWDPPERGAPESTFKGTPTAENPEYLGLDVYV 1255
 DB 1201 GGAAPQPHPPAFSPAFDNLVYWDPPERGAPESTFKGTPTAENPEYLGLDVYV 1255

RESULT 15
 ADE63281
 ID ADE63281 standard; protein; 1255 AA.
 XX
 AC ADE63281;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human Protein P04626, SEQ ID NO 9218.
 XX
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SN1; Chung.
 OS Homo sapiens.
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX

PF 14-AUG-2002; 2002WO-US025765.
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX WPI, 2003-268312/26.
 DR GENBANK; F04626.
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX Claim 1; Page; 1017p; English.
 PS
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNR)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1255 AA;
 Query Match 100.0%; Score 6815; DB 7; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 YNYLSTVSGCTLVCPILHNOEVTAEDETQRCCKSKFCARVCYGLGMEHLREVAVTSAN 360
 DB 301 YNYLSTVSGCTLVCPILHNOEVTAEDETQRCCKSKFCARVCYGLGMEHLREVAVTSAN 360
 QY 361 IOEFAGCKKIFGSLAFLPESFSDGPASNTAPLOEEOLOVFETLEITNGYLYISAMPBSLP 420
 DB 361 IOEFAGCKKIFGSLAFLPESFSDGPASNTAPLOEEOLOVFETLEITNGYLYISAMPBSLP 420
 QY 421 DLSVFONLQVIRGRILHNGAVSLTLOGLISWGLRLSRLBLSGLALIHNTHLCPVHTV 480
 DB 421 DLSVFONLQVIRGRILHNGAVSLTLOGLISWGLRLSRLBLSGLALIHNTHLCPVHTV 480
 QY 481 PMDOLFNNPQALHTANRPREDECVGBGLACHQICANGHCMBPPTQCVNSQRLRGEC 540
 DB 481 PMDOLFNNPQALHTANRPREDECVGBGLACHQICANGHCMBPPTQCVNSQRLRGEC 540
 QY 541 VEECRVLOGLPREVYNAHCLPCHPEGOPONGSVTCGPEADQCVACHYDPPFCVARC 600
 DB 541 VEECRVLOGLPREVYNAHCLPCHPEGOPONGSVTCGPEADQCVACHYDPPFCVARC 600
 QY 601 PSQVCPDLSYMPIMKFPDEBGACQPCPINCTHSCVDLDDKGCAPAEORASPLTISIAYVG 660
 DB 601 PSQVCPDLSYMPIMKFPDEBGACQPCPINCTHSCVDLDDKGCAPAEORASPLTISIAYVG 660
 QY 661 ILLVVLGVVFGIILIKRQOKIRKRYTMRRLIOETELVBLTPSGAMPNOAQMRILKTEL 720
 DB 661 ILLVVLGVVFGIILIKRQOKIRKRYTMRRLIOETELVBLTPSGAMPNOAQMRILKTEL 720
 QY 721 RKVAVLGSAGVYKGIWIPDGENVKIPVAIKYLRNTSPKAKKEILDEAYVWAGVSP 780
 DB 721 RKVAVLGSAGVYKGIWIPDGENVKIPVAIKYLRNTSPKAKKEILDEAYVWAGVSP 780
 QY 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHYENRGRISQDLNMQIAKMSLYEDVR 840
 DB 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHYENRGRISQDLNMQIAKMSLYEDVR 840
 QY 841 LVHBDLAARNVLVKSPNVKTTDGLARLDDIDETVHADGKPIKMMALSLIRRF 900
 DB 841 LVHBDLAARNVLVKSPNVKTTDGLARLDDIDETVHADGKPIKMMALSLIRRF 900
 QY 901 HQSWVSGYTVMLMTFGAKPYDGIPIAREIPDLLEKGERLPOPPICITDVMYIMVKCM 960
 DB 901 HQSWVSGYTVMLMTFGAKPYDGIPIAREIPDLLEKGERLPOPPICITDVMYIMVKCM 960
 QY 961 IDSECRPRFRELVEFSRMAADPQRFVVTQNEDELGPASPLDSTFYRSLLEDMDGLVDA 1020
 DB 961 IDSECRPRFRELVEFSRMAADPQRFVVTQNEDELGPASPLDSTFYRSLLEDMDGLVDA 1020
 QY 1021 EBYIVPQOGFFCPDPAGAGGMVHHRSSSTRSGGDLTLGLBSEBARSPPLABSEG 1080
 DB 1021 EBYIVPQOGFFCPDPAGAGGMVHHRSSSTRSGGDLTLGLBSEBARSPPLABSEG 1080
 QY 1081 AGSDVPFGDLSGMAKGLQSLPTHDPSPLOKYSDDPYVPLPSFSEIDGYVAPLTCSQPEXY 1140
 DB 1081 AGSDVPFGDLSGMAKGLQSLPTHDPSPLOKYSDDPYVPLPSFSEIDGYVAPLTCSQPEXY 1140
 QY 1141 NQPDVRPQPSPRSGPLPAARPAAGATLERPKTSLPGKNGVVDYFARGAVENBEYLTQ 1200
 DB 1141 NQPDVRPQPSPRSGPLPAARPAAGATLERPKTSLPGKNGVVDYFARGAVENBEYLTQ 1200
 QY 1201 GGAAPQPHPPAPSPADNLYYWDQDPPERGAPESTFGTPTAENBEYLGIDVY 1255
 DB 1201 GGAAPQPHPPAPSPADNLYYWDQDPPERGAPESTFGTPTAENBEYLGIDVY 1255

Search completed: January 3, 2006, 11:12:19
 Job time : 204 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2006, 11:05:47 ; Search time 50 Seconds
(without alignments)
2415.041 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815

Sequence: 1 MELALCRWGLLLALLPPCA.....TFKGTPTANPEYGLDVPV 1255

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6815	100.0	1255	1 A24571	protein-tyrosine k
2	5997	88.0	1260	1 TVRTNU	protein-tyrosine k
3	5993.5	87.9	1254	2 I48161	p-185 precursor
4	3169	46.5	1210	1 GQHUE	epidermal growth f
5	3145	46.1	1210	2 A53183	epidermal growth f
6	3122.5	45.8	1223	1 TVCHIV	epidermal growth f
7	3004.5	44.1	1308	2 A47253	epidermal growth f
8	2708	39.7	1166	1 S06142	protein-tyrosine k
9	2432.5	35.7	1342	2 A36223	kinase-related tra
10	2347.5	34.4	1339	2 JCA387	epidermal growth f
11	1765.5	25.9	698	1 TVFVLV	protein-tyrosine k
12	1702	25.0	604	1 TVYVUH	protein-tyrosine k
13	1653.5	24.3	1230	1 GQFPE	epidermal growth f
14	1645	24.1	544	2 S35745	protein-tyrosine k
15	1638	24.0	545	2 S00727	kinase-related tra
16	1621	23.8	540	2 B44776	protein-tyrosine k
17	1619	23.8	540	1 TVFVEB	protein-tyrosine k
18	1536	22.5	644	2 A36325	epidermal growth f
19	1301	19.1	1223	2 B88257	protein-tyrosine k
20	1301	19.1	1374	2 S70712	protein-tyrosine k
21	1211	17.8	1369	2 S70713	protein-tyrosine k
22	1177	17.3	1717	1 A45558	epidermal growth f
23	1155	16.9	527	2 A42032	epidermal growth f
24	997.5	14.6	843	2 A27131	epidermal growth f
25	814.5	12.0	346	2 S13807	protein-tyrosine k
26	757	11.1	311	2 S13808	protein-tyrosine k
27	736	10.8	1363	2 T43220	insulin-like growt
28	717	10.5	1382	1 INHUR	insulin receptor p
29	710	10.4	1383	2 A36080	insulin receptor p

ALIGNMENTS

RESULT 1
A24571
protein-tyrosine kinase (BC 2.7.1.112) erbB2 precursor - human
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein erbB
C:Species: Homo sapiens (man)
C>Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 05-Oct-2004
C:Accession: A24571; A25491; A44488; B44188; I59509; I57622
R:Yamamoto, T.; Ikawa, S.; Miyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;
Nure 319, 230-234, 1986
A>Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth f
A:Reference number: A24571; MUID:86118663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <YAM>
A:Cross-references: UNIPROT:P04626; UNIPARC:UPI000003F55F; GB:X03363; NID:g31197; PID:N
R:Semba, K.; Yamamoto, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A>Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epiderm
A:Reference number: A25491; MUID:86016729; PMID:2995967
A:Accession: A25491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: UNIPARC:UPI000016A0A7; GB:M11767; NID:G182163; PID:N:AAA35808.1; PID:
R:Cousens, L.; Yang-Peng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, P
Science 230, 1132-1139, 1985
A>Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromos
A:Reference number: A44188; MUID:86070181; PMID:2999974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COU>
A:Cross-references: UNIPARC:UPI000016AA26; GB:M12036; NID:G183988; PID:N:AAA35978.1; PID:
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALF', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: UNIPARC:UPI00001725C7; GB:M11730; NID:G183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A>Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 832-909 <REX>
A:Cross-references: UNIPARC:UPI0000070A3F; GB:L29395; NID:G459807; PID:N:AAA35809.1; PID:
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A>Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional
A:Reference number: I57622; MUID:87286698; PMID:3033551
A:Accession: I57622
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAU>

A:Cross-references: UNIPARC:UP1000000427; GB:M6792; NID:9183983; PIDN:AA58637.1; PID:
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
 C:Genetics:
 A:Gene: GDB:ERBB2; NCL; NEU; HER-2
 A:Cross-references: GDB:120613; OMIM:164870
 A:Map position: 17q21.1-17q21.1
 A:Introns: 25/1; 75/3; 147/1; 863/3
 A:Note: the list of introns is incomplete
 C:Function:
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
 inase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:122-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
 F:22-653/Domain: extracellular #status predicted <EXT>
 F:70-304/Domain: EGF receptor extracellular domain repeat <EE1>
 F:395-605/Domain: EGF receptor extracellular domain repeat <EE2>
 F:654-675/Domain: transmembrane #status predicted <TM>
 F:676-1255/Domain: intracellular #status predicted <INT>
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif
 F:68,124,181,229,530,571,629/Binding site: carbohydrate (Aen) (covalent) #status predict
 F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:753/Active site: Lys #status predicted
 F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 100.0%; Score 6815; DB 1; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 3.2e-277;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALACRWGLLALLPPGAATVCTGDMKRLPASPEHLDMLRLHYOCQVQGNL 60
 DB 1 METALACRWGLLALLPPGAATVCTGDMKRLPASPEHLDMLRLHYOCQVQGNL 60
 QY 61 ELYTLPINASLSFLDIOBVGSTVLIANQVRQVPLQRLRYRGTLFEDNYALAVLDNG 120
 DB 61 ELYTLPINASLSFLDIOBVGSTVLIANQVRQVPLQRLRYRGTLFEDNYALAVLDNG 120
 QY 121 DPLNNTTAVTASPGSLRELDRLSLTEILKGGVLIQRNPOLCYQTIIIMKDIFFHNNOLA 180
 DB 121 DPLNNTTAVTASPGSLRELDRLSLTEILKGGVLIQRNPOLCYQTIIIMKDIFFHNNOLA 180
 QY 181 LTLIDTNSRACHPCSPCKSGSCWSESSDQSLTRTVACAGCARCKGRLPTDCHBQC 240
 DB 181 LTLIDTNSRACHPCSPCKSGSCWSESSDQSLTRTVACAGCARCKGRLPTDCHBQC 240
 QY 241 AAGCTGPRHSDCLCLFHNSGICELHCPALVTYNTDTFESMPNDEGRYTTGASCVTACP 300
 DB 241 AAGCTGPRHSDCLCLFHNSGICELHCPALVTYNTDTFESMPNDEGRYTTGASCVTACP 300
 QY 301 YNYLSTDVSGCTLVPLRNLQEVTAEDGQRCKGKPCARVYGYGMHLEBRVAVTSAN 360
 DB 301 YNYLSTDVSGCTLVPLRNLQEVTAEDGQRCKGKPCARVYGYGMHLEBRVAVTSAN 360
 QY 361 IOEFAGCKKIFGSLAFLPESFPDPSANTAPLQEPQLQVFTLEETITGYLISAMPDLP 420
 DB 361 IOEFAGCKKIFGSLAFLPESFPDPSANTAPLQEPQLQVFTLEETITGYLISAMPDLP 420
 QY 421 DLSVFQNTQVTRIGRILHNGAVSLTQIGIGISWLGIRSLRELGSGIALIHNTHLCFVHTV 480
 DB 421 DLSVFQNTQVTRIGRILHNGAVSLTQIGIGISWLGIRSLRELGSGIALIHNTHLCFVHTV 480
 QY 481 PMDOLFRPHOALHTANRBECEVGBGLACHQLCARHGMGPGTLOCVNCOSQPLRGEC 540
 DB 481 PMDOLFRPHOALHTANRBECEVGBGLACHQLCARHGMGPGTLOCVNCOSQPLRGEC 540
 QY 541 VEECHVLQGLPREYVNAHCLCPHEPCOPONGSVTCFGEADQCYACAHYKDPFCVARC 600
 DB 541 VEECHVLQGLPREYVNAHCLCPHEPCOPONGSVTCFGEADQCYACAHYKDPFCVARC 600
 QY 601 PSGVPRDLSYMWIPKPEDEGACQPCPINCTHSQVLDLDDKGPAPQASPLTSTISAVVG 660
 DB 601 PSGVPRDLSYMWIPKPEDEGACQPCPINCTHSQVLDLDDKGPAPQASPLTSTISAVVG 660

DB 601 PSGVPRDLSYMWIPKPEDEGACQPCPINCTHSQVLDLDDKGPAPQASPLTSTISAVVG 660
 QY ILLVVLGVVGGILIKERQOKIRKXTMRRLDQETELVEPLTPSGAMPNOAMRLIKETEL 720
 DB 661 ILLVVLGVVGGILIKERQOKIRKXTMRRLDQETELVEPLTPSGAMPNOAMRLIKETEL 720
 QY 721 RKVVLGSGAFGVYKGIWIPDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAVGSP 780
 DB 721 RKVVLGSGAFGVYKGIWIPDGENVKI PVAIKYLRNTSPKANKELIDEAYVMAVGSP 780
 QY 781 YVSLRLGICLTSTVQLVTOQLMPVGCCLDHYENNGRLGSDOLMMQOIAAGMSYLEBYR 840
 DB 781 YVSLRLGICLTSTVQLVTOQLMPVGCCLDHYENNGRLGSDOLMMQOIAAGMSYLEBYR 840
 QY 841 LVHNDLAARNVLYKSPHNVKITTDFGLARLDDIDETEVHADGKVPKIMMALESILRRFT 900
 DB 841 LVHNDLAARNVLYKSPHNVKITTDFGLARLDDIDETEVHADGKVPKIMMALESILRRFT 900
 QY 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMTMVKCM 960
 DB 901 HQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYMTMVKCM 960
 QY 961 IDSECRPRFRELVSFBRMADPQRFVLTQEDIEDGASPLDSTYRSLLEDDMDGLVDA 1020
 DB 961 IDSECRPRFRELVSFBRMADPQRFVLTQEDIEDGASPLDSTYRSLLEDDMDGLVDA 1020
 QY 1021 EBYLVPOQGFPCPDPAAGAGVHRRSSSTRSGGDLTLGLPSEBEARSPPLAPSEG 1080
 DB 1021 EBYLVPOQGFPCPDPAAGAGVHRRSSSTRSGGDLTLGLPSEBEARSPPLAPSEG 1080
 QY 1081 AGSDVFDGLMGAKGLQSLPTHDPSPLQRYSDPVPPLPSEFDGYAPLTCSPOPEYV 1140
 DB 1081 AGSDVFDGLMGAKGLQSLPTHDPSPLQRYSDPVPPLPSEFDGYAPLTCSPOPEYV 1140
 QY 1141 NOPVRPOPSPBRGRLPARBAGATLBRPKTSLPGKNGVQDVAFAGCAVENBEYLTPQ 1200
 DB 1141 NOPVRPOPSPBRGRLPARBAGATLBRPKTSLPGKNGVQDVAFAGCAVENBEYLTPQ 1200
 QY 1201 GGAAPQHPHPPAFPAFDNLVYMDQDPPERGAPSTFGPTAENPEYLGIDVY 1255
 DB 1201 GGAAPQHPHPPAFPAFDNLVYMDQDPPERGAPSTFGPTAENPEYLGIDVY 1255

RESULT 2

TYR2NU

protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 05-Oct-2004

C:Accession: A24562; A61204

R:Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.

Nature 319, 226-230, 1986

A>Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.

A:Reference number: A24562; MUID:86118662; PMID:3945311

A:Accession: A24562

A:Molecule type: mRNA

A:Residues: 1-1260 <BAR>

A:Cross-references: UNIPROT:P06494; UNIPARC:UP10000161B83; EMBL:X03362; NID:956745; PIDN

R:Maui, T.; Mann, A.M.; Maratze, T.L.; Garland, B.M.; Okamura, T.; Smith, R.A.; Cohen,

Carcinogenesis 12, 1975-1978, 1991

A>Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no m

2-thiazolyl[1]formamide or N-methyl-N-nitrosourea.

A:Reference number: A61204; MUID:92035293; PMID:1682063

A:Accession: A61204

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 637-663, 'V', 665-702 <MS>

A:Cross-references: UNIPARC:UP100001725C8

A:Note: authors translated the codon GCA for residue 25 as Val

C:Genetics:

A:Gene: neu

C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology

C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>
 F:658-686/Domain: transmembrane #status predicted <TMN>
 F:723-988/Domain: protein kinase homology <KIN>
 F:731-739/Region: protein kinase ATP-binding motif
 F:71,791,263,535,576,634/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted
 F:758/Active site: Lys #status predicted
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 88.0%; Score 5997; DB 1; Length 1260;
 Best Local Similarity 87.9%; Pred. No. 4,3e-243;
 Matches 1105; Conservative 49; Mismatches 101; Indels 2; Gaps 2;

```

1  MELAALCRKGLLALLPFGAASVQVCTGDMKRLRPAPEETHLDMRLHYOGCQVVGQNL 60
4  MELAAKCRKMGFLALLPFGIACTQVCTGDMKRLRPAPEETHLDMRLHYOGCQVVGQNL 63
61  ELTYLPTNASSFLQDIEVQGYVLIANQVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
64  ELTYVPANASSFLQDIEVQGYVLIANQVQVPLQRLIRVGTQLFEDKXALAVLDNR 123
121  DPLANTTPVT-GASPEGIAELQRLSTELIKGGVLIQRNPQLCYQDTILMKDIFHKNQDL 179
124  DQDNVAASPTGRTPEGLRELQRLSTELIKGGVLIQRNPQLCYQDMVLKQVFRKNQDL 183
180  ATLIDITNRSRACHPCSPCKGSRCKGESSBDQSLTRTVACAGGACARCGPLPTDCCHQ 239
184  AYVDIDITNRSRACHPCSPCKGSRCKGESSBDQSLTRTVACAGGACARCGPLPTDCCHQ 243
240  CAAGCTGPRGSDCLACLFHNHSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTAC 299
244  CAAGCTGPRGSDCLACLFHNHSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTTC 303
300  PNYNSTDVGSCTLVCPPLNQEVTADGTQRCCKSKPCARVCYGIAMHLEAVNAVISA 359
304  PNYNSTEVSCTLVCPPLNQEVTADGTQRCCKSKPCARVCYGIAMHLEAVNAVISA 363
360  NIOEFGACKKIFGSLAFLEPESFGDPASTAPLQEPOLQVFETLSEITGYLIASAMPDL 419
364  NVOEFGACKKIFGSLAFLEPESFGDPASTAPLQEPOLQVFETLSEITGYLIASAMPDL 423
420  PDLVSFQNLQVIRGRILNNGAYSFLTQGLGISMVGLRSLRBLGSLALIHNTHLCEVNT 479
424  RLSLVSFQNLQVIRGRILNNGAYSFLTQGLGISMVGLRSLRBLGSLALIHNTHLCEVNT 483
480  VPMWDLFRPHQALHTANRPEDE-CVGEGLACHQICAGRCCKGPPPTCVNCSQPLRQ 538
484  VPMWDLFRPHQALHTANRPEDE-CVGEGLACHQICAGRCCKGPPPTCVNCSQPLRQ 543
539  ECVBERQVLTQGLPREVYNARHCLPCHREGOPONGSVTCGPREADOCVACAHYKDPFCVA 598
544  ECVBERQVLTQGLPREVYNARHCLPCHREGOPONGSVTCGPREADOCVACAHYKDPFCVA 603
599  RCPGSGVPLDLYMPPIKWPDEBEGACQPCINCTHSCVDLDDKCPAEQASPLTISIAV 658
604  RCPGSGVPLDLYMPPIKWPDEBEGACQPCINCTHSCVDLDDKCPAEQASPLTISIAV 663
659  VQILLVVLGVVFGIILKRRQKIRKTYMRLLQETELVEPLTPSGAMNQAOMRLKET 718
664  ECVLLFLILVVVVGILIKRRQKIRKTYMRLLQETELVEPLTPSGAMNQAOMRLKET 723
719  ELRKVKVLSGSGAFGTYYKGIWI PDGENYKI PVAIKYLRNTSPKANKELLDAYVMAAGV 778
724  ELRKVKVLSGSGAFGTYYKGIWI PDGENYKI PVAIKYLRNTSPKANKELLDAYVMAAGV 783
779  APYVSLRLGI CLTSIVQVLTQMLPYGCLLDHYRBNRGLSQDOLNMCQIAKMSYLED 838
784  SPYVSLRLGICLTSIVQVLTQMLPYGCLLDHYRBNRGLSQDOLNMCQIAKMSYLED 843
839  VALVHRDLAARVNLVKSPMHVKITDPLGLARLLDIDETRYHADGKRPITPMALLESILRR 898
844  VALVHRDLAARVNLVKSPMHVKITDPLGLARLLDIDETRYHADGKRPITPMALLESILRR 903

```

```

899  FTHQSDVMSYGVTVNELMTFGAKPYDGI PAREIPLLEKGERLPPPICTIDVYIMVYC 958
904  FTHQSDVMSYGVTVNELMTFGAKPYDGI PAREIPLLEKGERLPPPICTIDVYIMVYC 963
959  WMIDECRRPRELVSERSRNARDQRRVYVIONELGPAASPLDSTFFYSILLEDMDGDLV 1018
964  WMIDECRRPRELVSERSRNARDQRRVYVIONELGPAASPLDSTFFYSILLEDMDGDLV 1023
1019  DAEYLVPOQGFPCDPA PGAGGWVHHRRSSSTSGGDLTLGLEPSEEARSPPLAPS 1078
1024  DAEYLVPOQGFPCDPA PGAGGWVHHRRSSSTSGGDLTLGLEPSEEARSPPLAPS 1083
1079  EGAGSDVDFDGLGKMAAGLQSLPTHDPSPLOQRYSEDPVLPSEIDGVYALTCSPQPE 1138
1084  EGAGSDVDFDGLGKMAAGLQSLPTHDPSPLOQRYSEDPVLPSEIDGVYALTCSPQPE 1143
1139  YNQPDPVRPQPPSPREGPLPAARPGATLIERKTLSPGKNGVYKQVFAFGAVENPEYLT 1198
1144  YNQPDPVRPQPPSPREGPLPAARPGATLIERKTLSPGKNGVYKQVFAFGAVENPEYLT 1203
1199  PGGAAPQHPPEAPSPAFNDLYYWDODPERGAPSTFKGPTAEENPEYLGADVPV 1255
1204  PREGTASPPHSPAPSPAFNDLYYWDODPERGAPSTFKGPTAEENPEYLGADVPV 1260

```

RESULT 3

148161
 P-185 precursor - golden hamster
 C/Species: Mesocricetus auratus (golden hamster)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Oct-2004
 C/Accession: 148161
 R/Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa, Gene 140, 251-255, 1994
 A/Title: Cloning and activation of the Syrian hamster neu proto-oncogene.
 A/Reference number: 148161; M01D:94193007; PMID:7908275
 A/Accession: 148161
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-1254 <RES>
 A/Cross-references: UNIPROT:Q60553; UNIPARC:UP1000012A111; GB:D16295; NID:g493236; P1DN: C/Genetics:
 A/Gene: neu
 C/Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology
 C/Keywords: ATP
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif

Query Match 87.9%; Score 5993.5; DB 2; Length 1254;
 Best Local Similarity 87.7%; Pred. No. 5.9e-243;
 Matches 1101; Conservative 57; Mismatches 96; Indels 1; Gaps 1;

```

1  MELAALCRKGLLALLPFGAASVQVCTGDMKRLRPAPEETHLDMRLHYOGCQVVGQNL 60
1  MELAAKCRKMGFLALLPFGIACTQVCTGDMKRLRPAPEETHLDMRLHYOGCQVVGQNL 60
61  ELTYLPTNASSFLQDIEVQGYVLIANQVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
61  ELTYVPANASSFLQDIEVQGYVLIANQVQVPLQRLIRVGTQLFEDKXALAVLDNR 120
121  DPLANTTPVTGASPEGIAELQRLSTELIKGGVLIQRNPQLCYQDTILMKDIFHKNQDL 180
121  DPLANTTPVTGASPEGIAELQRLSTELIKGGVLIQRNPQLCYQDTILMKDIFHKNQDL 180
121  DPLANTTPVTGASPEGIAELQRLSTELIKGGVLIQRNPQLCYQDTILMKDIFHKNQDL 180
121  DPLANTTPVTGASPEGIAELQRLSTELIKGGVLIQRNPQLCYQDTILMKDIFHKNQDL 180
181  LTLIDITNRSRACHPCSPCKGSRCKGESSBDQSLTRTVACAGGACARCGPLPTDCCHQ 240
181  LTLIDITNRSRACHPCSPCKGSRCKGESSBDQSLTRTVACAGGACARCGPLPTDCCHQ 240
181  LTLIDITNRSRACHPCSPCKGSRCKGESSBDQSLTRTVACAGGACARCGPLPTDCCHQ 240
181  LTLIDITNRSRACHPCSPCKGSRCKGESSBDQSLTRTVACAGGACARCGPLPTDCCHQ 240
241  AAGCTGPRGSDCLACLFHNHSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTAC 300
241  AAGCTGPRGSDCLACLFHNHSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTTC 300
301  PNYNSTDVGSCTLVCPPLNQEVTADGTQRCCKSKPCARVCYGIAMHLEAVNAVISA 360

```



```

Db 301 YNYLSTEWGCTVCPVLPNNQEVTAEDGTQRCCKSKSCARVCYGLGMEHLRGARAIITSAN 360
QY 361 IOEPAGCKKIFGSLAFLPESFDGPDASNTAPLOPQOLQVFETLEITGTYLISAMPDILP 420
Db 361 IOEPAGCKKIFGSLAFLPESFDGPDASNTAPLOPQOLQVFETLEITGTYLISAMPDILH 420
QY 421 DLVSFQNTQVIRGRILLHNGAVSLTQGLGISWLGSLRLSRLGSLGLIHHNTHLCFVHTV 480
Db 421 DLVSFQNTQVIRGRVILHNGAVSLATQGLGISWLGSLRLSRLGSLGLVLIHRTNTHLCFVHTV 480
QY 461 PMDQLFRRPHQALLHTANRPEDECGEGLAGHQQLCARHGWMPGRTQCVNCSQELRGQEC 540
Db 461 PMDQLFRRPHQALLHTSNGPSEECGLKDFACYPCLAHGCMWPGPTQCVCNCSHFLRGQEC 540
QY 541 VEECHVLQGLPREYVNAARHCLFCHPECOPONGSVTCFGEADQCAACHVADPPPCVARC 600
Db 541 VKECHVWGLPREYVNGKHLFCHEECOPONSTETCTGSEADQCAACHVADPPPCVARC 600
QY 601 PSQVPRDLSYMPIMKYPDEBGMCOCPINCTHSCVDLDRGCGPAEORASPATSIATVVG 660
Db 601 PSQVPRDLSYMPIMKYPDEBGMCOCPINCTHSCVDLDRGCGPAEORASPATSIATVVG 660
QY 661 ILVVVVLGVNPGIILKRRQOKIRKTYTMRLLQETLVEPLTPSGAMPNOAQMRILKETEL 720
Db 661 ILFLVIGVVGILLKRRQOKIRKTYTMRLLQETLVEPLTPSGAMPNOAQMRILKETEL 720
QY 721 RKVKVLGSGAFQTVYKGIWIPDGEVNVKIPIVKYVRENTSPKANBEILDEAVYMAVGSP 780
Db 721 RKVKVLGSGAFQTVYKGIWIPDGEVNVKIPIVKYVRENTSPKANBEILDEAVYMAVGSP 780
QY 781 YVSRLLGICLTSTVLTQMLPYGCLLDHVRNRRGLSODILLNQCWIJAKMSYLEDVR 840
Db 781 YVSRLLGICLTSTVLTQMLPYGCLLDHVRNRRGLSODILLNQCWIJAKMSYLEDVR 840
QY 841 LVHRLAARNLVKSPNHVKITDFTGLARLDDTEHYHADGSKVPIKMTALESIIRRF 900
Db 841 LVHRLAARNLVKSPNHVKITDFTGLARLDDTEHYHADGSKVPIKMTALESIIRRF 900
QY 901 HOSDWMSGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYMIWVKCM 960
Db 901 HOSDWMSGVTVWELMTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYMIWVKCM 960
QY 961 IDSECRPRRELIVSEFSRMAARDPQRFVVIQNEIDLPASPLDSTFYRSLLEDMDGLVDA 1020
Db 961 IDSECRPRRELIVSEFSRMAARDPQRFVVIQNEIDLPASPLDSTFYRSLLEDMDGLVDA 1020
QY 1021 BEYLVPOGQGFCEPDAPAGAGVHHRHSSSTRSGGDLTGLRESEEAAPRSLABSEG 1080
Db 1021 BEYLVPOGQGFCEPDAPAGAGVHHRHSSSTRSGGDLTGLRESEEAAPRSLABSEG 1080
QY 1081 AGSDVFDGDLGMAKGLQSLPTHPSPLOQYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140
Db 1081 AGSDVFDGDLGMAKGLQSLPTHPSPLOQYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140
QY 1141 NQPDVROPSPREBEPDPAARAGATLERPKTLSPGKGVVVDVAFGCAVENPEYLTQ 1200
Db 1141 NQPDVROPSPREBEPDPAARAGATLERPKTLSPGKGVVVDVAFGCAVENPEYLTQ 1200
QY 1201 GGAAQOPHPAPAFSAPFNLYWDDPPERKGPSTFKGTPTPAENPEYLGIDVPY 1255
Db 1201 GGAAQOPHPAPAFSAPFNLYWDDPPERKGPSTFKGTPTPAENPEYLGIDVPY 1255
QY 1254 GGSASQPH-PPALCPAFDMLYWDQPSERGSPPNTFQGTPTPAENPEYLGIDVPY 1254
Db 1254 GGSASQPH-PPALCPAFDMLYWDQPSERGSPPNTFQGTPTPAENPEYLGIDVPY 1254

```

RESULT 4

GQHUR
 epidermal growth factor receptor precursor - human
 N>Contactins: protein-tyrosine kinase (EC 2.7.1.112) erbb
 C:Species: Homo sapiens (man)
 C:Date: 15-Nov-1984 #sequence revision 27-Nov-1985 #text change 05-Oct-2004
 C:Accession: A00641; A25772; S30024; A36872; A43615; A23062; A05281; A60143; A33
 R:Ulrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Y
 ig, P.H.
 Nature 309, 418-425, 1984

A>Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression of
 A:Reference number: A00641; MUID:84219729; PMID:6328312
 A:Accession: A00641
 A:Molecule type: mRNA
 A:Residues: 1-1210 <L1N>
 A:Cross-references: UNIPROT:P00533; UNIPARC:UPI0000050F30; EMBL:X00588; NID:g31113; PIDN
 A>Note: the authors translated the codon AAG for residue 540 as Asn
 R:Rishi, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.
 Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985
 A>Title: Characterization and sequence of the promoter region of the human epidermal gro
 A:Reference number: A25772; MUID:85270438; PMID:2991899
 A:Accession: A25772
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-29 <ISH>
 A:Cross-references: UNIPARC:UPI000016A882; GB:M1234; NID:g181981; PIDN:AAA2370.1; PID:
 R:Haley, J.; Whittle, N.; Bennett, P.; Kinchington, D.; Ulrich, A.; Waterfield, M.
 Oncogene Res. 1, 375-396, 1987
 A>Title: The human EGF receptor gene: structure of the 110 kb locus and identification o
 A:Reference number: S30024; MUID:88217333; PMID:3329716
 A:Accession: S30024
 A:Molecule type: DNA
 A:Residues: 1-29 <HA2>
 A:Cross-references: UNIPARC:UPI000016A882; EMBL:X06370; NID:g31118; PIDN:CAA29668.1; PID
 R:Haley, J.D.; Waterfield, M.D.
 J. Biol. Chem. 266, 1746-1753, 1991
 A>Title: Contributory effects of de Novo transcription and premature transcript terminat
 A:Reference number: A36872; MUID:91107677; PMID:1988448
 A:Accession: A36872
 A:Molecule type: DNA
 A:Residues: 1-29 <HAL>
 A:Cross-references: UNIPARC:UPI000016A882; GB:M38425; NID:g181977; PIDN:AAA63171.1; PID:
 R:Xu, Y.; Iehli, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.; Mer
 Nature 309, 806-810, 1984
 A>Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs
 A:Reference number: A00642; MUID:84245835; PMID:6330563
 A:Accession: A00642
 A:Molecule type: mRNA
 A:Residues: 150-187, 'KSVIOAV', 195, 'W', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-321
 A:Residues: 713-964 <LIN>
 A:Cross-references: UNIPARC:UPI00001725BE
 A:Experimental source: carcinoma cell line A431-7
 R:Lin, C.R.; Chen, W.S.; Krulliger, W.; Stolarczyk, L.S.; Weber, W.; Evans, R.M.; Verma, I.
 Science 224, 843-848, 1984
 A>Title: Expression cloning of human EGF receptor complementary DNA: gene amplification
 A:Reference number: A43615; MUID:84196372; PMID:6326261
 A:Accession: A43615
 A:Molecule type: mRNA
 A:Residues: 713-964 <LIN>
 A:Cross-references: UNIPARC:UPI00001725BE
 A:Experimental source: epidermal carcinoma cell line A431
 R:Stamen, F.A.; Gope, M.L.; Schulz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.
 Biochem. Biophys. Res. Commun. 124, 125-132, 1984
 A:Reference number: A23062; MUID:85046483; PMID:6093780
 A:Accession: A23062
 A:Molecule type: mRNA
 A:Residues: 1028-1210 <STM>
 A:Cross-references: UNIPARC:UPI00001725BF
 R:Weber, W.; Gull, G.N.; Speiser, J.
 Science 224, 294-297, 1984
 A:Reference number: A05281; MUID:84172183; PMID:6324343
 A:Accession: A05281
 A:Molecule type: protein
 A:Residues: 25-30, 'S', 32-51, 454-467 <WEB>
 A:Cross-references: UNIPARC:UPI00001725CO; UNIPARC:UPI00001725CI
 R:Russo, M.W.; Lukas, T.J.; Cohen, S.; Scarso, J.V.
 J. Biol. Chem. 260, 5205-5208, 1985
 A>Title: Identification of residues in the nucleotide binding site of the epidermal grow
 A:Reference number: A60143; MUID:85182650; PMID:2985580
 A:Accession: A60143
 A:Molecule type: protein
 A:Residues: 740-744, 'X', 746-747 <RUS>

A:Cross-references: UNIPARC:UPI00001725C2
 R:McCzokewi, B.; Mosig, G.; Cohen, S.
 Nature 309, 270-273, 1984
 A>Title: ATP-stimulated interaction between epidermal growth factor receptor and superec
 A:Reference number: A38023; MUID:8419154; PMID:6325948
 A:Contents: annotation; receptor activity
 A:Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA
 R:Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.
 Cell 59, 33-43, 1989
 A>Title: Functional independence of the epidermal growth factor receptor from a domain
 A:Reference number: A33331; MUID:90003233; PMID:2790960
 A:Contents: annotation; internalization signal
 C:Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor c
 C:Geneid6:
 A:Gene: GDB:EGFR
 A:Cross-references: GDB:120610; OMT:131550
 A:Map position: 7p12.3-7p12.1
 C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
 F:1-24/Domain: signal sequence #status predicted <Sig>
 F:25-1210/Product: EGF receptor #status predicted <NMT>
 F:25-645/Domain: extracellular #status predicted <EXT>
 F:75-300/Domain: EGF receptor extracellular domain repeat <EE1>
 F:390-600/Domain: EGF receptor extracellular domain repeat <EE2>
 F:646-668/Domain: transmembrane #status predicted <TM>
 F:669-1210/Domain: intracellular #status predicted <INT>
 F:710-975/Domain: protein kinase homology <KIN>
 F:718-726/Region: protein kinase ATP-binding motif
 F:999-1046/Region: coated-pit mediated internalization signal
 F:1047-1210/Region: inhibitory
 F:128-175,352,413,444,528,603/Binding site: carbohydrate (Asn) (covalent) #status predic
 F:745/Active site: Lys #status experimental

Query Match 46.5%; Score 3169; DB 1; Length 1210;
 Beet Local Smlarity 49.8%; Pred. No. 3,8e-125;
 Matches 632; Conservative 177; Mismatches 345; Indels 116; Gaps 22;

11 LLLALLPREGA--STQVCGTDMKRLRPAPEPTHLDMLRHLYOGGQVQGNLELYPTN 68
 14 LLAALCPASRALEBKVKVCGTGNKLTQLCTFEDHFLSLDRMNCEVVLGNLEITYVQRN 73
 69 ASLSFLODIOEVQGVYLLAHNOVQVPLQRLRIVRGTOLEBDNYALAVLNDGPIANTTP 128
 74 YDLSPLKTIQEVAGVYLLALNTVERIPLENLQIRGNMYENSYALAVLNDG----- 126
 129 VTGASPGGIREPLQSLSLTEILKGVLIQRNPOLCTQDITLIMKDIHKNQLALTLIDTKR 188
 127 ---ANKTGILKEIPMBRLQRIHGAVERFSNNPALCVNESIQMRDIVSSDPLSNMSDPQNH 183
 189 SRAICPSPMCKSGSCWSSSEDCSLPTTVCAAGCA-RCKGRLPTDCCHEGCAAGCTGP 247
 184 LGSCKCKDSCPNCGSGWAGEENCOCLYKICAOQSGCRKSPDCCNCCAAAGCTGP 243
 248 KHSDCIACIHPNHSICELHCPALVTYNTDTESMPNPEGRATYFGASCVTACPYNYLSTD 307
 244 RSDSLVCKKPKDEATCKOTCPRLMYNTTYQMDVNPBGKXSFATYCKKPRRYVYTD 303
 308 VGSCTLVCPRLHNOVTAEDGTORCEKSPCARVCYGLAMEHLRRAVATSNANIOBPAC 367
 304 HGSVCRAACADSYEM-EEBDGVAKCKKCEGPRCKVCGIGIGFQKSLISINATNIKGFQK 362
 368 KIFPSLAFPSPOGDPAISNAPLQBPOLQVPELTLEETLVLYLSAWDLSLDLSVPRN 427
 363 TSISSDLHLPLVAFPGDSTFHTPRDLPOLDLTKVTKETLIGVLLIQAMENNTDLHAFEN 422
 428 LOVIRGRILHNGAVSLTGLGLISWLGSLRELISGLALLHHNTHLCFVHTVPMDOJFR 487
 423 LEIIRGRTQKHOFSLAVSLVLTSLGLSLKEISGDVYIIGNKMLCANITNNKKLFG 482
 488 NPHQALLTANRPEDECVESGLAQHLCARGHCWGPPTQVNCQOFLRGQCEVEECRYL 547
 483 TSGQTKTILSNRGENSKATQGVCHALCSPEGCWGPBPDPDVSCHNVSGREGRCVCKKLL 542

548 QGLPREYVNAHRCPLRCHPECOFONGSVTCFGBEADQVACAHYKDPPECVACPSGVKPD 607
 543 EGEPRPEFVNSEFCIQHCECLPQAMNINCTFGRGPNDCIQCAHYIDHPCVCTKCPAGVNGE 602
 608 LSYPIKRPPEBEGACQPCPINCHISCYVDLDDKGPACQASPLTSIISAVG---ILLY 664
 603 NNTL-VMKYADAGHCHLCHPCTYGCQFGLGEGPTNGPKP-SITGVMGALLLLV 659
 665 VLVGAFGLIKRROQKIRKXTMRLLQETELVEPLTPSGAMPNOAIRIKETELRKVK 724
 660 VALGIG---LPMRRRHIVKRTLRLLDLRELVEPLTSGAPNOLAIRIKETEPKIK 716
 725 VLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRNTSPRANKIIDEAVYMGVSGPYVR 784
 717 VLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRNTSPRANKIIDEAVYMGVSGPYVR 776
 765 LIGITLSTVQLVQLMAYGCLLDVRENRKRLSGODLNNCMQAKMSYLEVRLVHR 844
 777 LIGITLSTVQLVQLMAYGCLLDVRENRKRLSGODLNNCMQAKMSYLEVRLVHR 836
 845 DLAAENVLKSNNHYKINDPGLARLLDITETVYADGGKVPKMMALESIARRPPTHOSD 904
 837 DLAAENVLKTQHVKTIDPGLAKLLGAEKRYAAGKVPKMMALESIARRPPTHOSD 896
 905 VMSYGVTVWELMTFGAKPYDGIPAEIPDLLEKGRRLPQPICTIDVYIMVKKMMIDSE 964
 897 VMSYGVTVWELMTFGSKKYDGIPIASEISILEKGRRLPQPICTIDVYIMVKKMMIDSE 956
 965 CRPRFRELVSERMAARDPQRFVVIQ-NEGLQASPLDSTFYRSLLEDNDKGLVDABEY 1023
 957 SRPKFRELIIEFSKNAARDPQRYVLVQGERNMLPSPTSNFRLMDEEDMDVDVABEY 1016
 1024 LVPOGFPCPRDPAAGGVNHHNRSSSTGGGDLTLGLEPSEEARSPAPLASEGAGS 1083
 1017 LIPQGGF-----SSGSTRTPLSLSLSTS 1042
 1084 DVFDGLMGAKGLQSLPTDPSPLQRYJEDPTVPLPSET--DGYVVALTCSPOREYVN 1141
 1043 N--NSTVACIDRNGQSPICEDSFLQRYSDPTALNEDSIDDFL-----PVPEYIN 1094
 1142 QPDVAPPPSPREGPLPAAPRAGATL---ERPKTSLPGKNGVVKVPAFAGAVENBEY 1196
 1095 Q-----SVP-KRPAQSVQNPVYHNQPLNAPSPSDPHYD--PHSTAGNBEY 1138
 1197 L-TPQGAAPQHPPAPSPAPDNLYWDO-----DP-----PERGAPSTPKG 1239
 1139 LNTVQ-----PTCVNSTFDSPAHWAQKSHOISLDNPYQODFPPEKAKPNGIKFG 1189
 1240 TPTAENPEYL 1249
 1190 S-TAENAEYL 1198

RESULT 5
 A3183
 epidermal growth factor receptor precursor mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 05-Oct-2004
 C:Accession: A3183; A43818; S24942; A28941; S45325; I49643
 R:Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Baty, H.S.; Jenkins, N.A.;
 Genes Dev. 8, 399-413, 1994
 A>Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor t
 A:Reference number: A3183; MUID:94170986; PMID:8125255
 A:Accession: A3183
 A:Molecule type: mRNA
 A:Residues: 1-1210 <LDB>
 A:Cross-references: UNIPROT:Q01279; UNIPARC:UPI0000175614; GB:U03425
 R:Aviv, A.; Lax, I.; Ullrich, A.; Schlesinger, J.; Givol, D.; Morae, B.
 Oncogene 6, 673-676, 1991
 A>Title: Comparison of EGF receptor sequences as a guide to study the ligand binding sit
 A:Reference number: A43818; MUID:91232866; PMID:2030916
 A:Accession: A43818
 A:Molecule type: mRNA

A:Residues: 1-714 <AVI>
A:Cross-references: UNIPARC:UPI0000175615; GB:X59698
R:Eisinger, D.P.; Sertero, G.
submitted to the EMBL Data Library, June 1992
A:Reference number: S24942
A:Accession: S24942
A:Molecule type: mRNA
A:Residues: 969-971, 'K', 973-1115, 'D' <EIS>
A:Cross-references: UNIPARC:UPI0000175616; EMBL:Z12608
R:Heisermann, G.J.; Gill, G.N.
J. Biol. Chem. 263, 13152-13158, 1988
A:Title: Epidermal growth factor receptor chreonine and serine residues phosphorylated
A:Reference number: A28941; MUID:8830814; PMID:3138233
A:Accession: A28941
A:Molecule type: protein
A:Residues: 689-694, 'X', 696-704, 'L', 706-707, 989-992, 'XX', 995-996, 'X', 998-1000, 1002-1009,
A:Cross-references: UNIPARC:UPI0000175617; UNIPARC:UPI0000175618; UNIPARC:UPI0000175619;
R:Hibbs, M.L.; Dunn, A.R.; Alexander, W.S.
submitted to the EMBL Data Library, April 1994
A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor
A:Reference number: S45325
A:Accession: S45325
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-971, 'K', 973-1210 <VER>
A:Cross-references: UNIPARC:UPI00002182B; EMBL:X78987; NID:9488830; PIDN:CA55587.1; PI
R:Paria, B.C.; Das, S.K.; Andrews, G.K.; Dey, S.K.
Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993
A:Title: Expression of the epidermal growth factor receptor gene is regulated in mouse b
A:Reference number: 149643; MUID:93126380; PMID:7678348
A:Accession: 149643
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 12-20, 22-132 <RES>
A:Cross-references: UNIPARC:UPI000016C026; GB:L06864; NID:9193001; PIDN:AA53029.1; PID:
A:Gene: EGFR
C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology
C:Keywords: ATP, growth factor receptor; kinase-related transforming protein; phosphoprc
P:1-24/Domain: signal sequence #status predicted <SIG>
P:646-670/Domain: transmembrane #status predicted <TM>
P:712-977/Domain: protein kinase #status predicted <KIN>
P:720-728/Region: protein kinase ATP-binding motif
R:696,699/Binding site: phosphate (Thr) (covalent) #status experimental
R:697,1070,1071/Binding site: phosphate (Ser) (covalent) #status experimental
R:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental
P:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental
P:1157/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 46.1%; Score 3145; DB 2; Length 1210;
Best Local Similarity 49.7%; Pred. No. 3.8e-124;
Matches 635; Conservative 169; Mismatches 353; Indels 120; Gaps 24;

OY 11 LLLALLPFGAA--STOYCTGTDMLRLPASPTLHMLRLHLYOGCQVQGNIELTYLPTN 68
DB 14 LLLTALCAAGALAEKXKVCSTSNRLTOLGTFEDHFLSLQRMYNCEVVLGNLEITYVORN 73
OY 69 ASLSFLDIOIEQVGYVLAHNOVROVPLORLIVYGTOLFEDNYALAVLDNGDPLNNTTP 128
DB 74 YDLSLKTIQEVAGVYVLAHNTVERIPLENLDIIRGNALYENTYVALILSN----- 124
OY 129 VTGASPGGLRELQLSLTEILKGVLIQRNPOLCYODTILMKDI-----FHKNQIALTLI 184
DB 125 -YGTNRTGRLRLPMNLQELILGAVRFSNRPILCMMDTIQWRDIQVNFMSMSMDL--- 180
OY 185 DTNRSRACHPCSPMKGSRCESSSDCQSLTRTYCAGGCA-RCKGRLPLTDCHEQCAAG 243
DB 181 -QSHPSSCPCKDCPSCPNCSCWGSGEENCOKLTKIICAAQCCSHRCGRSPSDCHNQCAAG 239
OY 244 CTGPHSGTCLALHTNHSGLICLHCPALVTYNTDTPESNPNEGRTYFGASCVTACPNY 303
DB 240 CTGPRPSDCLVQKQFQDEATCKOTCPRLMLNPTTYQMDVNPBGKISFGATCVKCKPRNY 299

OY 304 LSTDVSGCTVCPRLHNOEVTAEADSTORCEKSKPCARVCYGLGMEHLREVAVTSANTIOE 363
DB 300 VVDHSGCVRAKCGPDYEV--EEDGIRKCKCDKCGCRKCNIGIGIEGFDTSINATNIKH 358
OY 364 FAGCKTIFGSLATLPESFDGPRASNTAPLOEQOVETLEITGYLYISAMPDPLSL 423
DB 359 FKCYTASGDIHLIPVAFKDSFTPTPLDLRELILKTVETLIVETLITQAMPDMDTLH 418
OY 424 VFQNIQVIRGRILHNGAVSLTLQGLISWLGARSRLREKSGSLALIHNTHLCPFYTPWD 483
DB 419 AFENLEIRRGTRKHGQESLAVGLNITSLGRSLKELSDSDVYISGRNLCYANTINMK 478
OY 484 QLPENPHQALHTANRPDECEVSGGLACHQICANHCSCGPEPTCCVNCSCQLRQCEYEE 543
DB 479 KLFTPNQTKIKMNRARAKDCKAVNHCNPLCSSEGCSPGPRPCVSCQVSRGECVEK 538
OY 544 CRVLQGLPREVYNAHCLPCHPECOQPNQSVTCTCGPRADOCVACAHYKDPFCVARGCSG 603
DB 539 CNLIEGEPRFVENSECTIQCHPECLPQAMNITCTGRGPDNCTIQCAHYIDGPHCVKTCFAG 598
OY 604 VKPDLVYPIKFPDEEGACQPCPINCTSCVDLDDKGCSPAEORASPLTISIISAVGILL 663
DB 599 IMGEMNTL-VKTKADANNVCHLCHANCTYCGAGGLQCEVWPSGPKIPSLATIGVGLL 657
OY 664 VVLGVVFGI-LIRROQKIRKTYMRRLLOETELVEPLTPSGAMPNOQMRLKTELARK 722
DB 658 FIVV-VALGIGLPMRRRHIVAKTRLRLQRELEVEPLTPSGEAPNQAHILILKETERPK 716
OY 723 VKVLGSGAFGVYGVIMIPGSENVKIPVALIVLEENSPKANKILDEAYVMAGVGSRYV 782
DB 717 IKVLGSGAFGVYGLMIPSEKVKIPALIKELBAISPKANKILDEAYMASVDNDHV 776
OY 783 SRLIGICTSTVQLVTOQLMPYGCILLDVHRENRGRLSGODLLNMCQIAGKSYLEDVRLV 842
DB 777 CRLIGICTSTVQVLTQLMPYGCILLDVHRENRGRLSGODLLNMCQIAGKSYLEDVRLV 836
OY 843 HRDLAARNVTVKSPNNHYKITDFGLARLLIDDETYHADGKVPPIKMALESILRRPFTHQ 902
DB 837 HRDLAARNVTVKTPQHVKITDFGLAKLLGABEKEYHAEKGKVPPIKMALESILRRPFTHQ 896
OY 903 SDVWSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGRLEPPICTIDVVMIMYKCMID 962
DB 897 SDVWSYGVTVWELMTFGSKPYDGI PASDISLEIKERLEPPICTIDVVMIMYKCMID 956
OY 963 SECPREFELVSEFSRMAPQRFVVIQ-NEDLCPASPLSTFYRSLLEDMDMDVDAE 1021
DB 957 ADSRPFRELILERSQMARQRLVYIQGDSRMHLPSFTDSNFYRALMDEDMEDVVDAD 1016
OY 1022 EYLVPOQGFPCPDPAFGAGKVVHHRSSSTRSGGDLITGLPSEBEAPRSPPLAPSGA 1081
DB 1017 EYLVPOQGF-----NSPST-----SRTPLLSLISA 1042
OY 1082 GSDVFDGDLGMAKGLQSLPETHDSPLOQTSSEDPVLPSET--DGVAVALTCSPOPEY 1139
DB 1043 TSN-----NSTVACINRNGSCRVKEDAFQRYSSPFTGALVEDNIDDAVL-----PVPEY 1092
OY 1140 VNQPDVAPQPPSPREGRPLPARPAGATL-----BRPTLSFGKGVADVAFGGAENP 1194
DB 1093 VNQ-----SVF-KRPAGSVQNPVYHNGRPLHAPRGLRHTQN--PHNAGVNP 1136
OY 1195 EYL-TPQGAAPQHPPPAPSPAFDNLVYWDQ-----DP-----BERGAPSTF 1237
DB 1137 EYLVTAQ-----PTCLSSGNSPALMTQKSHQMSLDNPYQODFPFKETKPNIGIF 1187
OY 1238 KGTPTAENPEYLGIVP 1254
DB 1188 KG-PTAENAEYLRVAPP 1203

RESULT 6
TYCHIV
epidermal growth factor receptor precursor - chicken
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbb

C:\Species: Gallus gallus (chicken)
C:\Date: 28-Feb-1986 #sequence_revision 05-May-1995 #text_change 05-Oct-2004
C:\Accession: A27720; A00643
R:\Lax, I.; Johnson, A.; Hawk, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Vennet
Mol. Cell. Biol. 8, 1970-1978, 1988
A:\Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in mou
A:\Reference number: A27720, MUID:88261272; PMID:3260329
A:\Accession: A27720
A:\Molecule type: mRNA
A:\Residues: 1-1223 <LAX>
A:\Cross-references: UNIPROT:P00534; UNIPARC:UPI00001725C3; GB:M20386
R:\Nielsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rotman, F.M.; Crittenden, L.B.; Raines, M
Cell 41, 719-726, 1985
A:\Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and pro
A:\Reference number: A00643; MUID:85228222; PMID:2988784
A:\Accession: A00643
A:\Molecule type: mRNA
A:\Residues: 585-1223 <NTL>
A:\Cross-references: UNIPARC:UPI00001725C4; GB:M10066
C:\Genetics:
A:\Gene: erbB
C:\Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology
C:\Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor
specific protein kinase
F:\1-30/Domain: signal sequence #status predicted <SIG>
F:\31-1223/Product: epidermal growth factor receptor #status predicted <MAT>
F:\31-654/Domain: extracellular #status predicted <EXT>
F:\81-307/Domain: EGF receptor extracellular domain repeat <EE1>
F:\397-610/Domain: EGF receptor extracellular domain repeat <EE2>
F:\655-677/Domain: transmembrane #status predicted <TM>
F:\678-1223/Domain: intracellular #status predicted <INT>
F:\719-984/Domain: protein kinase homology <KIN>
F:\727-735/Region: protein kinase ATP-binding motif
F:\136-202,280,351,370,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent) #
F:\197-650/Binding site: carbohydrate (Ser) (covalent) #status predicted
F:\687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:\758/Active site: Lys #status predicted
F:\1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Query Match 45.84; Score 3122.5; DB 1; Length 1223;
Best Local Similarity 48.74; Pred. No. 3,36-123;
Matches 632; Conservative 175; Mismatches 345; Indels 145; Gaps 25;

8 RMGLLALIPPGA-----STVCTGTDMKLRLPASBETHLMDLRHLYOGCQVVGSL 61
Db RGAALVLLLLGLVALGSAVEKKVCGGTNNKLTQLGHEDHPTSLQRMYNCEVLSNLE 72
QY LTYLPTNASLSFLQDIOEVQGYVLIANNORQVPLQRLRYRGTLFEDNVALVLDNGD 121
Db LTYVEHNRDLPTFKTIOEVAGYVLIANNVDVPLENQITRGVNLVYNSFALAVLSNYH 132
QY PLANNTPVVGASGGGRELDRSLTEILKGSVLIQRNPOLCYQDTIIMKDIHKNQNL 181
Db LKQ-----GRELPMKRLSETILNGSVKLSNNPKLCMDIVLNMDDITDSK-PL 182
QY TLID-TNRSRACHPCSPMCKSGSRCESSBEDQSILRTVCAGAGCA-RCKGPLPTDCHEQ 239
Db TYLDKPSNLSGCCKCHNCETEDHCGAGEQNCOTILTKVICAQCGSGRCRGKVPSCCHN 242
QY CAAGCTGPKHSDCIACLHPHSGICELHCPALVYTNDDTFESMNPBGRYTFGASCVTAC 299
Db CAAGCTGPRESDCLACRKFRRDDATCKDTCPLVLYNPTTYQMDVAPBEGYSGATCVRE 302
QY PNYVLTSDVSGCLTVGLCHNHQEVTAEDGTORCEKSKRCACVACVGLGMEHLREVAVTA 359
Db PNYVLTSDHSCSVRSCNTDTYEV-BENGVRCKCKKDCGLCSKVCKNGIGELGILISINAT 361
QY NTQEPAGCKKIFGSLAFLPSFPGDPASNTAPLOPEQLQVETLEBITYLYISAMPDSL 419
Db NIDSFNCKTINDDVSLTPVAFGLDAFTKTLPLDQKLDVFRYVKEISGFLIIQAMPDVA 421
QY PDLVSPQNTQVIRGRILHNGAYSLTQGLGISWGLGSLRLREGLALIHNTHLCSFVAT 479

422 IDLVAFENLEIIRGRTRKHGGYSLAVLVNLTQISGLRSLKEISDDDIAMKKNKLCYADT 481

480 VPMIDLFNNPQIALHTANRPEDCEVGEGLCHQICAGHCWGGRPTQVCNCSQTLRQE 539

482 MAMWSLFATOSQKTKITONRNKNDCTADRHVCDPLCSGVGCGWGPAPFHCFSGRFSSRKE 541

540 CVEECRLVQGI.PREYVNAARHCLPCHPRECQPNONG---SVTCGFEPADOCVACHYDPPFC 596

542 CVKQCNITIQGSPREBERDSKCLPCHSECLVONSTAYNTTGGPDPDHCKAHFLDGHHC 601

597 VAPCSGVKPLSYMPRIWKFPDEBEGACOPCINCTHSCVDDDKGCPAEORASPLTISIIS 656

602 VKACPGATLGENDTL-VKAYADANAACOLCHPNCTRGCKGGEBCP---NGSKTSPIAA 657

657 AVV-GILLVVVLGVVFGILIRROOKIRKTYMRRLLOETELVEPLTPSGAMPNOAKRIL 715

658 GVGGGLCTLVVVGIGIGLYLRRR-HIVAKRTLRRLLORELVEPLTPSGEAPNOAHLRIL 716

716 KETELRKVTUGSAGFGVTVKGIWIPDSENVKIPAIYVLEBNSPKNKAIILEAYYMA 775

717 KETEKKKVKGSSGFGTVYGLMIPSEKVKIPALTELEASVPRKNKIILEAYYMA 776

776 GVGSPPVYSRLIGICLTSTVQLVTOLMPYGCILLDHVRENGRLSGODLLMCMQIAKAGNSY 835

777 SVDPNPHVCRLLGICLTSTVQLITOLMPYGCILLDYIREKNQIGSYLLNMCVQIAKAGNSY 836

836 LEDVALVHRDLAARNVLKVSBNHYKIDTGLARLLIDETEHYADGKVPRIKMALESIL 895

837 LEERLTVHRDLAARNVLKVPQHYKITDFGLAKLIGADEKEYHAGGKVPRIKMALESIL 896

896 RRRFPHGSDWVSQYTVTWELMTFGAKPYDGI.PABEIPDLLEKGRILPQPICTIDVYIM 955

897 HRIYTHOSDWSYGTVMELMTFGSKPYDGI.PABEISSVLEKGERLPQPICTIDVYIM 956

956 VKCWMIDSECRPREPELSEFSRMAADPQIFVVIC-NEDLCPASPLDSTFYRSLLEDMD 1014

957 VKCMNIDDSKPRELILAEFSKKAADPPKILVIOGDNRMLPSTBDKFTRLMEEDM 1016

1015 GDLYDAEBEYLVPOOGFFCPDPAFGAGVNHRRHSSSTRSGGDLTLGLBPSBEERPS 1074

1017 EDIYDADEYLVPHQOFF-----NSPST-----SRT 1042

1075 L-----ABSEAGSDVFPFGDLGMAKAGLQLPHNDSPFLQRYSDPVLPLSET--DGY 1127

1043 LLSSLSAITSNNASATNCTID-----RNGQCHPVRBDSFVORYSSDPGNFLBESIDG 1094

1128 VAPLTCSQPRPVYVQNPQVRPQSPREBGRPLAARPAAGATLRLPKTILSPGKNGVVDV-- 1185

1095 L-----PAPRYVNO--LMPKPS-----TANVQNIYNNISLT 1125

1186 -----AFGAVENPEYLTPOGDAPOPHPPAPSPAVDNLVYWDQ----- 1225

1126 AISKLPMSRKQONHSTAVDNPYL-----NTNGSPPLAKTVPESSPVMIQSNHGIN 1177

1226 -DPPE-----RGAPSTFGKPTPAENPEYLTGLDVP 1254

1178 LDNPYQODFLPNETKPMGLKVPAAENPEYLTVAAP 1214

RESULT 7

A47253

Epidermal growth factor receptor, HER4 - human

C.Species: Homo sapiens (man)

C.Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 05-Oct-2004

C.Accession: A47253

R.PIowman, G.D.; Clouascou, J.M.; Whitney, G.S.; Green, J.M.; Catlton, G.W.; Foy, L.; N

Proc. Natl. Acad. Sci. U.S.A. 90, 1746-1750, 1993

A.Title: Ligand-specific activation of HER4/p180erbB4, a fourth member of the epidermal

A.Reference number: A47253; MUID:93189574; PMID:8383326

A.Statut: preliminary; not compared with conceptual translation

A.Molecule type: nucleic acid

A.Residues: 1-1308 <P/O>

A:Cross-references: UNIPROT:015303; UNIPARC:UP100000499DF; GB:107868; NID:9337359; PIDN:
 A:Note: sequence extracted from NCBI backbone (NCBI:126842)
 C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology
 C:Keywords: ATP; growth factor receptor
 F:716-981/Domain: protein kinase homology <KIN>
 F:724-732/Region: protein kinase ATP-binding motif

Query Match 44.1%; Score 3004.5; DB 2; Length 1308;
 Best Local Similarity 45.6%; Pred. No. 36-118;
 Matches 614; Conservative 183; Mismatches 375; Indels 175; Gaps 27;

```

QY 9 WGLLALLPFGAA-----STQVCTGDMKRLRLPASEPETHLDMRLHYOGCOVQGNLELY 64
DB 8 WWWWLVAAGTVQPSDSQSCVAGTENKLSLSDEQYRALRKYYENCEVVMGLIETS 67
QY 65 LPTNLSFLDIQIOGVVLLIAHNOVRQVPLRLRIYRGTOLEFDNYALAVLDGDDPLN 124
DB 68 IEHRDLSTLRVREVTGYVVALNQFRLPLENRIIRGTLYEDRLALAIPLNRYKDG 127
QY 125 NTTPTVGASPGGLRELOSLRLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNQALATLI 184
DB 128 NF-----GLQELGLNLTLEILNGGVYVDNQKFLCYADTIIHMODIVANPWSNLTLV 178
QY 185 DTRNRACHPCSPCKSGSCWSESDQSLRTTYCAGGC-ARCKGRLPTCCCHQCAAG 243
DB 179 STNGSSGGGRCHKSCG-RCWGPTEHNCQTLRTYCARQDCRCYGPVSDCCHRECAAG 237
QY 244 CTGPHGSDCLACLFHNSGICELHCPALVTYNTDFEESWNPBEGRTYGASCVTACPNY 303
DB 238 CSGPQDTQCFACMNNDSGACTQCPQTFVVPPTTFQLEHNNAKITGAFVCKKCPNF 297
QY 304 LSTDVSGCTLVCPHNOEVTAEADGTQRCCKSPCARVCYGLGMEHLREVAVTSANTIQE 363
DB 298 V-VDSSSCVRACPSPSKMEY-EENGIMCKRPTDIPCXACDGGTSLMSAQVDSNIDK 355
QY 364 FAGCKITGSLAFLEBSPDGPASNTAPLOPQLOVFELEITGYLYISAMPDLPDS 423
DB 356 FINCKINGNMLFLVTGTHGDPYNAIEAIDPEKLVNFRVREITGFLNIQSPPMPTFS 415
QY 424 VFQNLQVRGRILHNGAVSLTLOGIGISWLGSRISRELSGSLALHNHTLCEVYTPMD 483
DB 416 VPSNLVTIGKRVLYGSLILIKQOIGTSLQOSLKEISAGNITYTDSNLCYTTIMWT 475
QY 484 QLFNPNQALHTANRPDECEVGEGLACHQLCARGHCWGPPTQVNCSSQFLRGEVBE 543
DB 476 TLFSTINQRIYVIRDRKKAENCTAEQMVNCHLCSGCGPDPQCLSCRRFSGRICIES 535
QY 544 CRVLQGLREYVNAHCLPCHPEQCP-QNGSVTCGPRPADQVCAAHKKDPFCVAPRPS 602
DB 536 CNLYDGERREFENGSIQVECDPQCKMEDGLTTCGPGPDNCTKCSHFKDGNCVEKCPD 595
QY 603 GVKPDLSTMPILWKFPEDEGACQCPINCTHSCVDLDDKGC-----PAEQRAPL 651
DB 596 GLQGNSTF--IFKPADPRBECHPCNPCTOGNGFTSHDICYYPWTGHTLPQHR----- 649
QY 652 TTSISAVV--GILLVVLGVVFGILIKRQOKIRRYTMRLLQETLEVEPLTPSGAMPNQ 709
DB 650 TPLIAGVIGGLFILIYVIGLFAVYVRRRSIK-KKRALRRFL-ETELVEPLTPSGAIDNQ 707
QY 710 AQMLTKETELKRVNVLGSGAGTYKGIWIPDGENVYIPIVAKILRENTSKAKKELTD 769
DB 708 AQRLTKETELKRVNVLGSGAGTYKGIWIPDGENVYIPIVAKILRENTSKAKKELTD 767
QY 770 EAYVWAGSPVSVSLGICLTSTVQLVTQMLPYCCLDLHVENEGRLSGODLAWMCQI 829
DB 768 EALIMASDHRLVALLGVCLSPITQVTQMLPHGCLLEYHENCDNIGSOLLAWMCQI 827
QY 830 AKGMSYLEDVRLVHDLAARNVLYKSPNHVKITDGLARLLDIDETEXHADGKVPKIM 889
DB 828 AKGMVYLEERRLVHDLAARNVLYKSPNHVKITDGLARLLDIDETEXHADGKVPKIM 887
QY 890 ALESLTRRFTQSGVSVGYTVWELMTFGAKPYGCIARRELPDLLEGERLPPPICTI 949
DB 890 ALESLTRRFTQSGVSVGYTVWELMTFGAKPYGCIARRELPDLLEGERLPPPICTI 949

```

```

DB 888 ALECIHYRKFTHQSDVMSYGVTIMELTGCKPYDGIPTREIPLDLKEGERLPPPICTI 947
QY 950 DVYIMVWCMWIDSECRPREFELVSEFSRMAADPQRFVIOQED-LGPASPLDSTFYRSL 1008
DB 948 DVYIMVWCMWIDSDSRKFKELAEFSRMAADPQRFVIOQEDDMKLPSPYDSKFPQNL 1007
QY 1009 LEDDDMGDLVAEYLYVQCGFCPPDPAPAGCMVHHRRSSSTRSGGDLTLGLEPSEE 1068
DB 1008 LDEEDLEMDMAEEYLYV-QAFNIPPP-----IYTSRARDISNMS-----EIGSPPPA 1055
QY 1069 EAPRS-----PLAP-SEGAGSVFPDGLGMAKGLQ 1099
DB 1056 YTPMSGNQFVRDGGFAEGGVSVYRAPRTIIPAPVAQHTAIFDSCNGTLKRV 1115
QY 1100 SLPTHPDPLQRYSEDPVPLPS-----ETDGYVALTCSPOPEYVNPQDVRQPPSP 1152
DB 1116 APHYGSDSTQRYSDPVPAPERSPRGELBEGYMTMRDQKQEYLNPNV----- 1167
QY 1153 RBGLPAPAPGATLERPKTSLPGKNGVAVQVAFAGVAVENPEYLTPQGAAPQHPPPA 1212
DB 1168 -ENPFVSRH-----KNGDLQ-----ALDPEYHNASNG-----PPVA 1198
QY 1213 -----FSPAFDNLVYWDQDPPERGA--PPST 1236
DB 1199 EDEYVNEPLYNTFANTLGKAEYLNKNIILSMPEKAKKAFDNDYWNHSLPPRSTLQHPDY 1258
QY 1237 FKGTPT-----AENPEYL 1249
DB 1259 LQEVSTKVFYKONGRIRPIVAENPEYL 1285

```

RESULT 8

S06142
 protein-tyrosine kinase (EC 2.7.1.112) mrk-γ precursor - southern platyfish
 N:Alternate names: epidermal growth factor receptor homolog; kinase-related transforming
 C:Species: Xiphophorus maculatus (southern platyfish)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-Oct-2004
 C:Accession: S06142; S13809
 R:Witbrodt, J.; Adam, D.; Maltzschek, B.; Maenele, W.; Raulf, F.; Telling, A.; Roberts
 Nature 341, 415-421, 1989
 A:Title: Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu loc
 A:Reference number: S06142; MUID:90015140; PMID:2797166
 A:Accession: S06142
 A:Molecule type: DNA
 A:Residues: 1-1166 <WTT>
 A:Cross-references: UNIPROT:P13388; UNIPARC:UP100001725C5; EMBL:X16891; NID:965290; PIDN
 R:Adam, D.; Maenele, W.; Scharf, M.
 Oncogene 6, 73-80, 1991
 A:Title: Transcriptional activation of the melanoma inducing Xmrk oncogene in Xiphophoru
 A:Reference number: S13807; MUID:91125882; PMID:1846957
 A:Accession: S13809
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 821-1025 'N', 1027-1098 'A', 1100-1166 <ADA>
 A:Cross-references: UNIPARC:UP100001715E2; EMBL:X56319; NID:965284; PIDN:CAA39763.1; PID
 C:Genetics:
 A:Gene: mrk
 A:Map position: Y
 A:Introns: 872/3; 898/1; 947/1; 979/3; 1025/3; 1056/1
 C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology
 C:Keywords: ATP; growth factor receptor; phosphotransferase; transmembrane protein; tyro
 F:1-25/Domain: signal sequence #status predicted <Sig>
 F:26-166/Product: kinase-related transforming protein (Tu) #status predicted <Mat>
 F:707-972/Domain: protein kinase homology <KIN>
 F:715-723/Region: protein kinase ATP-binding motif

Query Match 39.7%; Score 2708; DB 1; Length 1166;
 Best Local Similarity 45.5%; Pred. No. 6.2e-106;
 Matches 576; Conservative 164; Mismatches 389; Indels 138; Gaps 28;

```

QY 4 AALCRWGLLALLPFGAAST-----QVCTGDMKRLRLPASEPETHLDMRLHYOGCOVQGN 59
DB 8 AALLQ--LLLVLSISRCSTDPDRKVCQGTSONMTM---LDNHLYLKKMKMSGCVNVLN 62

```

```

60  LELTVLPNTASLSPLODIOEVQGVYLIANNOVQVPLQRLIRVGRNOLPEDNYALAVLDN 119
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
63  LEITTOENQDLSFQSIQIEVGIVLIANNEVSTIPLVNLRLIRGNLTIEGFTLLVMSN 122
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 120  GDPINNTTPVTGASPGELREQLRSITELIKGVLIOBNPOLCYDPTILMKDIFHKNNQL 179
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 123  YQK-NPSSP--DVYQGLKQLQSLNLTIELSGGVKSHNPLLCANETIMWMDIVDKTSP 179
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 160  ALTLIDITNRSRACHPCSPKCKSRCKGSESDCOSLTRVCGAGC-ARKGGLPDDCCH 238
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 160  TNNLLPHAFEROCOCODHCVCNASCWAPGPHCQKFTKLLCAEQCNRRRGRPKPDDCCE 239
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 239  OCAAGCTGPKSDCLACHFNHSGICELHCPALVTNTDTPESMPNREGRYTFGASCVTA 298
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 240  HCAAGCTGPRATDCLACRDFNDGTCOTCPPEKTIYDISHOVVDNPNKITFGACVKE 299
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 299  CPYNYLSTDVGSCTLVCPLANOEVTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTS 358
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 300  CPSNVVYVE-GACVRSACAGMLEVD-ENGKSKCKPCDGVCPKVCDDIGIGLSNTLAVNS 357
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 359  ANIOFPAGCKTIFGSLATLPESFDGDDPASNTAPLOEQLQVETIETTYIYISAMPDS 418
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 358  TNIRSFNCTKINGDIILNRNSPEGDPHYKIGTMDPEHLMNLTVCYKELTYVIMWBDN 417
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 419  LPDLSVPNLQVIRGRILHNGAYS-LTLQGLISWLGRLSRLSLGSLALIHNNHLCPV 477
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 418  MTSLSVPNLQVIRGRITTFSGFSFVVQVNRHLOWMLSLSLVSAVANILKNTLQLKRA 477
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 478  HTVPMDQFLRNPHQALLHTANRPEDECVGEGLACHQLCARHGMSPGPTOCVNSQFLRG 537
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 478  NTIMNRRLPRSEDOIEVDART-----ENQTCNNECSEDDCM-PGPTMCVSCILHVRG 529
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 538  QCBCEBCEVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEBADQVACAHAHYKDPKPV 597
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 530  GRCVASCNMLQGEPRQAQVGRCVQCHOECLVQTSITCYGPGPANCSKSAFQDGPQCI 589
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 598  ARCPGVPRLSYVMIKPRDEEGACOPRINCHSCVDLDDKGPAGQASPLSIISA 657
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 590  PRCPHIGLIDGDTL-IMKRADMGQOPHONCTGCCSGPGLSGRGD-IYSHSLAVGL 647
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 658  VVGLILVVVLGVFGILIKRQOKIRKYTMRLQETELVEPLTFPSGAMPNOQMRLIKE 717
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 648  VSGLLITVIALLVLLVLRRRRIK-RKRTIRCLLDEKELVEPLTFPSGAPNQAFLRIKE 706
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 718  TELRKVKVLGSGAFGTVYKGIWIPGENVKIPVAIKVIRENTSPYANKIILDEAYVMAV 777
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 707  TEFKDRVVLGSGAFGTVYKGLWNPDEGEMIRIPVAIKVIREATSPKYNQBEVLDEAYVMAV 766
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 778  GSPVYSLIGTCLSTVQLVYQMLPYGCLDHYRENBRGLSGODLLNMOQIAKMSYLE 837
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 767  DHPHVCRLGLCLTSAVOLVYQMLPYGCLDHYRHOERIICQOMLLNMCVOJAKMNYLE 826
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 838  DVRLVHRDLAARVNLVKSPPNHVKTIDFGIARLLIDETEHADGGKVPKMMALESIIR 897
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 827  ERHLVHRDLAARVNLVKSPPNHVKTIDFGIARLLIDETEHADGGKVPKMMALESIIR 886
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 898  RPTHOSDWSYGVYTWELMTFGAKPYDGIIPAREIPDLLEKGRLLPQPPICITDVMIMVK 957
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 887  TYTHOSDWSYGVYTWELMTFGSKPYDGIIPAKEIASVLENGERLLPQPPICITEVYMIILK 946
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 958  CMMTIDSEBRPRELVESESRMARDPQRFVIVONEDLPASPLDSTFYRSLIEDDDMDL 1017
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 947  CMMTIDSEBRPRELVESESRMARDPQRFVIVONEDLPASPLDSTFYRSLIEDDDMDL 1001
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1018  VDAEELVLPQOGFPCSPDPAPGAGVNHHRSSSTRSGGDLTLGLLEPSSEERAPSPAP 1077
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1002  VDADLYLPRYKRI-----NRQGS-----EPCIP 1024
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1078  SEGAGSDVFDGDLGMAKGLQSLPTHDSPLOQRYSEDPV-PLPSEPDGYVAPLTCSPQ 1136
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1025  PTGH-----PVRENSITLRNISDPTQNALREKLDLGH----- 1055
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

QY 1137  PEYVNPDPVRPQP-----PSPRE-----GLP-AARPAAGATLRRPKTLSPKNGVVD 1183
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1056  -EYVNPQPSSESSRSRLSDIYNPEYEDLTQGWGVSLSSQEAETNFRPELTANNQSL---- 1111
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1184  VFAFGAVENPREYLTTPQGAAPQHPPPAFSPAFNLYYWDODPPERGAAPSTFGTPTPA 1243
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1112  PLVSSGSMDDPDY---QAG-----YQAAAF-----LPQTALTGNGMFLPAA 1149
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1244  ENPEYLG 1250
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1150  ENLEYLG 1156
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
A36223
kinase-related transforming protein (erbB3) (EC 2.7.1.-) precursor - human
C:Species: Homo sapiens (man)
C:Date: 04-Oct-1991 #sequence_revision 13-Jan-1993 #text_change 31-Dec-2004
C:Accession: A36223; 159164
R:Kraus, M.H.; Issing, W.; Miki, T.; Popescu, N.C.; Aaronson, S.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 9193-9197, 1989
A:Title: Isolation and characterization of ERBB3, a third member of the ERBB/epidermal g
A:Reference number: A36223; MUID:90083234; PMID:2687875
A:Accession: A36223
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1342 <RBA>
A:Cross-References: UNIPROT:P21860; UNIPARC:UPI000017A3AE; GB:M29366
R:Plowman, G.D.; Whitley, G.S.; Neuhauer, M.G.; Green, J.M.; McDonald, V.L.; Todaro, G.J
Proc. Natl. Acad. Sci. U.S.A. 87, 4905-4909, 1990
A:Title: Molecular cloning and expression of another epidermal growth factor receptor-re
A:Reference number: 159164; MUID:90311312; PMID:2164210
A:Accession: 159164
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-559,'G',561-957,'P',959-1063,'G',1065-1342 <RES>
A:Cross-References: UNIPARC:UPI000050P2D; GB:M34309; NID:9183990; PID:AAA5979.1; PID
A:Genetics:
A:Gene: GDB:ERBB3; HER3
A:Cross-References: GDB:119860; OMIM:190151
A:Map position: 12q13-12q13
A:Keywords: ATP; phosphotransferase
F:707-972/Domain: protein kinase homology <Kin>
F:715-723/Region: protein kinase ATP-binding motif

Query Match 35.7%; Score 2432.5; DB 2; Length 1342;
Best Local Similarity 40.7%; Pred. No. 2.2e-94;
Matches 533; Conservative 191; Mismatches 458; Indels 129; Gaps 32;

QY 10  GILLALALPPGA--STOVCSTGDMKLRLPASPETHLDMRLHLYOGCOVVOGRLVLTLP 67
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 11  GILSLANGSEVNSQAVCPGLTNGLSVTDANEOYQLTKLYRCEVVMGRLVLTGH 70
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 68  NASISFLDIOEVQGVYLIANNOVQVPLQRLIRVGRNOLPEDNYALVLDNGDLNNTT 127
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 71  NADISFLQWIEVGVYLVANNEFSTLPLPMLRVVRGQYVDKRAIFVM-----LNTNT 125
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 128  PYTGASPGELREQLRSITELIKGVLIOBNPOLCYDPTILMKDIFHKNNQLALITDN 187
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 126  ----NSSHALQRLTLQTEILSGGVYIEKDKCMQDTIMRDLVRDRD--AEIVYKD 178
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 188  RSRACHPSPKCKSRCKGSESDCOSLTRVCGAGC-ARKGGLPPTCCCHQCAAGCTG 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 179  NGRSCPCHAYCKG-RCKGPGSEDDQTLTKITICAPQCNHGFGRPNPNCCHDECAAGGSG 237
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 247  PKHSDCLACHFNHSGICELHCPALVTNTDTPESMPNREGRYTFGASCVTAAPYNYLST 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 238  PQDPTCFACRHPFNHSGACVPRCPQGLVYNNKLTFQLEBPHTKYQYGVGVASCPHNFV-V 296
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 307  DVGSCITLVCPLANOEVTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSANIQEPAG 366
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 297  DQTSVCRACPPDKMEVD-KNGIKMCEPCGGLCPKACBEGTGG--SRFQTVSSNIDGFPV 353
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

367 CKKIFGSLAFLPESPDGDPASNTAPLOEOLQFETLEETITGYLIYSAMPDSDLPLSVFQ 426
364 CTKIIGNDLFLITGNGDPWHKIPALDPEKLNVTFTVEITGYLNIQSPMPHMFVS 413
427 NLQVIRGILHNGAVS-LTLQGLGISWIGLRSLRSLGSLALIHNTHLCPHTVYPMQL 485
414 NLTTGGSLVNRGSLIMKLNVTSLGFRSLKESIGRILYISNRQLCHHSLSMTKV 473
486 FRNPHQALLHTA-NRPEDECVGEGLAGHQLCARHGCMWPPTQCVNCSQFLRGQCEVEEC 544
474 LRGPTEERLDIKHNRPRRDCVABEGKCDPLSSGGCMWGPQCCLSCTNYSRGVCTVHC 533
545 RVLOGLPREYVNAHRLCPHPECOPONGSVTCFGEADQVACAHYKDPFPCVACPSGV 604
534 NFLNGEPREFAHEACFSCHPECQMEGATCNCGSGSDTCAQCAHFRDGPCHVSSCPHGV 593
605 KPDLSTMYIKRPDEGACQPCPINCTHSCTVDLDKGPABERA----SPLTISAVVG 660
594 LG-AKGPITYKPDVQNECRPCHEMCTGCGKPELDCLGQTLVLIGKTHLTMALTVIAG 651
661 ILVVVLGVFGILIKRQOKIR-KYTMRRLLQETELVEPLTPSGAMPNOAMRIKETE 719
652 --LVIFPMMLGTFLYMNGRRIQNRAMRRTYERGESIEPLDPS-EKANKVLARIKETE 708
720 LRKVVILSGAFGTVYKGIWIPDGENVKIPVAIKVLRNTSPKANKELIDEAYVNAVGS 779
780 PYSVSLIGTCLTSTVQVLTQMLPYGCLLDHVENNGRLGSDLLMWCQIWKMSYLEDV 839
769 AHIVVLCLCPRESSIQTLVQYPLSLSDHVNQHGAIGPOLLMWGVQIAGMYLIEH 828
840 RLVRHDLAARNVLVSPNHVKITIDFGLARLLDIDETEYHAGCKVPIKMALESILRRF 899
829 GMYHNNLAARVNLKSPSQOVADFGVADILRPDQKOLLYSAKTRIFMMALSHIFPKY 888
900 THQSDVMSYGVTVWELMTFGAKPYDGIAPAREIPDLLEKGERLPORPTCTIDVYMIWVKW 959
889 THQSDVMSYGVTVWELMTFGAEPYAGRLAEVPRDLLEKGERLAPQICTIDVYMWVKW 948
960 MIDSCRPRRELVSERFARMADPORFVYIOMEDIGRA---SPLDSTYRSLLEDDMGD 1016
949 MIDENIRTEKELANEFTRMARDPPRYALIKES-GPGIAPGEPHGLTNKKLEVELEP 1007
1017 LVDAEYLVLPQGFPCPDPAFGAGMVHHRSSSTREGGGDLITGLEP-SEEAERAPSL 1075
1008 ELDDLDLLEABED-----NLATTTLSALSLPVTGLNRPRESSQL 1048
1076 APSEGAGSDVDPDGLGMAKGLQSLPTHD-PSPLQRYSEDPVLP-----SETDGYV 1128
1049 SPSGGY-MPMQGNIGESCQESAVSGSSBRCPVSLH-----PMPRGCLASESSEGHV 1101
1129 A-----PLTCSPOPE-----YVNOGDVAPQPPSPREGP-----L 1157
1102 TGSEAELOEKVSMCRSRSRSPRPRGDSAYHSQHSLLTPVTPISPGLEBEDVNGYVM 1161
1158 PAARPAAGATLBRPKTISP-GKNGVY-----KDVAFGAAYENPEYLPPOGGAAPQHP 1210
1162 PDTHTAKGTPSSREGTSLSSVGLSVCTEBEDD-----EYVYNNRRRRHSP-PHP 1212
1211 PAFSPAPNLVYWD-----QDPERGAPSPSTFGKTPTAENPEYL 1249
1213 RPSSELELGEYEMDVGDLSLSTGSCPLHPVIMPAGTTPDEDEYEM 1263

RESULT 10
JC4387
epidermal growth factor receptor homolog precursor - rat
N:Alternate names: ErbB3 protein; HER3 protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Jan-1996 #sequence_reviation 19-Apr-1996 #text_change 05-Oct-2004
C:Accession: JC4387
R:Heilayer, N.J.; Kim, H.H.; Greaves, C.H.; Sierke, S.L.; Koland, J.G.

Gene 165, 279-284, 1995
A>Title: Cloning of the rat ErbB3 cDNA and characterization of the recombinant protein.
A:Reference number: JC4387; MUID:96096535; PMID:8522190
A:Accession: JC4387
A:Molecule type: mRNA
A:Residues: 1-1339 <HEL>
A:Cross-reference: UNIPARC:UPI000017A3DA; GB:U29339; NID:9915389; PID:9915390
A:Experimental source: liver
A>Note: The authors translated the codon AAC for residue 369 as Thr and GTT for residue
C:Comment: This protein is a functional heregulin receptor that transduces signals to th
C:Genetics:
A:Gene: ErbB3
C:Keywords: ATP; growth factor receptor; liver; phosphoprotein; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:120-1339/Product: epidermal growth factor homolog #status predicted <MAT>
F:640-659/Domain: transmembrane #status predicted <TMM>
F:705-970/Domain: protein kinase homology <KIN>
F:713-721/Region: protein kinase ATP-binding motif
F:939,1051,1156,1194,1196,1219,1257,1259,1273,1286,1325/Binding site: phosphate (Tyr) (C
Query Match 34.4%; Score 2347.5; DB 2; Length 1339;
Best Local Similarity 40.8%; Pred. No. 7,8e-91;
Matches 524; Conservative 170; Mismatches 434; Indels 155; Gaps 34;
3 LAALCRMGLLIALIPRGA--STQVCTGIDMKRLPASPEPTHLDMLHYQCGOVQGN 59
7 LQVLC---FLSLARGSEMNSQAVCPGTINGLSVTDADNOYQTLVLYKEKEVWGN 62
60 LELVYLPNTNASTLQDIOEYGVYVLAHNOVRQVPLQRLRYVGTOLFEYVYALVADN 119
63 LEIVLTGNADLSLQWIREYAYVLAANESVPLPMLNVVGTQVYDQKFAIVH-- 120
120 GDPVNTTTPVYGAAPGGLRELQLRSLTEILKGVLIQRNPQCYQDTILMKDIFKNNQL 179
121 ---LNVNT---NSHMLRQLKFTQLTEILSGVYIEKNDLCHMDITDMEDIVRV-- 170
180 ALTLIDITNRSPACHPCSPMCKSGSRCKGESSDCCSLRTVCAAGC-AACKPLRPDCCHE 238
171 GAEIVVKNNGANCPRCHEVCNG-RCWGPGPDCCOILTCTICAPQCNCRCPFPNPQCHD 229
239 QCAAGCTGPKKSDCLACLFHNSGICELHCPALVTYNTDPESMPNPRGRTPFGASCTYA 298
230 ECAGGCGSPQDTDFACRFRPDSACVPRCEPLVYNTKLTQLQLEBNPTTKIQYGGVCAVS 289
299 CPYNYLSTDVSGCTVLCVPLHNOEYTABDGTQRCCKSPCARVYCYGLMEHLREYVATS 358
290 CPHNFV-VDQFCYRACPRDMEVD-KHGLKMGCPGGLCPKACBGTSG--SRVQTVDS 345
359 ANIQEPACKKIFGSLAFLPESPDGDPASNTAPLOEOLQFETLEETITGYLIYSAMPDS 418
346 SNIGFVNTCTKIENLDFLITGLNVDPWHKIPALDPEKLNVTFTVEITGYLNIQSWBPB 405
419 LPDLSVQNLQVIRGILHNGAVS-LTLQGLGISWIGLRSLRSLGSLALIHNTHLCPV 477
406 MENSVSFNSLTTTGGSLVNRGSLIMKLNVTSLGFRSLKESIGRILYISNRQLCHHSLSMTKV 465
478 HTVPMQDLFRNPHQALLHTA-NRPEDECVGEGLAGHQLCARHGCMWPPTQCVNCSQFLR 536
466 HSLMNTRLRLRPSBERLDIKYDRPLGECIABGVKCDPLSSGGCMWGPQCCLSCTNYSR 525
537 GQCEYBECRVLOGLPREYVNAHRLCPHPECOPONGSVTCFGEADQVACAHYKDPFPC 596
526 EGVCVTHCNFQGGPREFVHNAQCFSCHPFCLPMEGISTYNGSSSDACARCAHFRDGPB 585
597 VAPCPGVKPDLSMPRIKRPDEGACQPCPINCTHS--VDLDKXGCPAQRASPLTSI 654
586 VNSCPHGILG--ANGPIYKIPDAQNECRPCHEMCTGCGKPELDCLGQAEVLSKPIV 643
655 ISAVGILLVVLGVFGILIKRQOKIR-KYTMRRLLQETELVEPLTPSGAMPNOAMR 713
644 IAVYVG--LNVILMILGSPLYMNGRRIQNRAMRRTYERGESIEPLDPS-EKANKVLAR 700
714 ILKETLRKVVILSGAFGTVYKGIWIPDGENVKIPVAIKVLRNTSPKANKELIDEAYV 773


```

Db      701  I F K E T L R K L K V L G S G V G T V H N K G I W I E G E S I K I P V C I K T I E D K S G S Q F O A V T D H L A      760
Qy      774  M A G V G S P R V S R L L G I C L T S T V Q L V T Q L M P R Y G C L L D H R E N G R L G S O L L M W C M Q I A G M      833
Db      761  V G S I D N H A I V A L L G I C P G S S I O L V T Q U P L G S L D H V K Q H R E T L G P O L L M W G V O I A G M      820
Qy      834  S Y L E D V R L V H R D L A A R N V L V K S P N H V K I T D F G L A R L D I D E T E Y H A D G K V P I K M A L E S      893
Db      821  Y L L E S H S V H R D L A R N V M L K S P S Q V A D F G V A D L L P R P D K O L L H S A K T P I K M A L E S      880
Qy      894  I L R R R F T H O S D V M S Y G V T V M E L M T F G A K P Y D G I P A R E I P D L L E K G E R L P O R P I C T I D V M      953
Db      881  I H P K Y T H Q S D V M S Y G V T V M E L M T F G A P R Y A G L R A E I P D L L E K G E R L A Q O I C T I D V M      940
Qy      954  I M V K C M I D E S C R P R F E L V S F S R M A R D P O R F V T I Q N E D I G P A S P L D S T Y R S I L E B D D      1013
Db      941  V M V K C M I D E N I R P T F K E L A N E F T M A R D P R Y L V I K R A S - G P G T P - P A A E P S V L T T K E      997
Qy      1014  M G D L V D A E E Y L V P R O G F R C P R P A R A G A G M V H N H R S S T R S G G D L T L G L P S E E - - - -      1068
Db      998  L - - - - - O E A L E P E L - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -      1023
Qy      1069  - - - - - E A P R S P L A P S E G - - - - - A G S D V P R D G L G M G A K G L S L P T H D      1105
Db      1024  L G S A L S L P T G L T R P R G S Q S L S P S S G Y M P N Q S S L G R A C L D S A V L G G R E Q P R S I S L H -      1082
Qy      1106  P S P L Q R Y S E D P T V L P S E T D G V - - - - - A P L - - - - - T C - - - - - S P O E - - - - - Y N Q P D V      1145
Db      1083  P I P R G R - - - - - P A S E S S E G H V T S E A E L Q B K V S V C S R S R S R S R P R D G S A Y H S Q R H S      1135
Qy      1146  R P O R P S P R E S P - - - - - L P A R A P A G A T L E R P K T L S P - G K N G V V - - - - - K D V A F      1187
Db      1136  L L T P T P L S P G L E E D G N G V M P D T L R G A S S R E G T L S S V G L S V L G T E B E D - - - - -      1191.
Qy      1188  G C A V E N P E Y L T P O G G A P O P H R P      1210
Db      1192  - - - - - E E Y E Y M N R K R G S P - P R P      1209

RESULT 11
TVFVLY
protein-tyrosine kinase (EC 2.7.1.112) erbB - avian leukosis virus
N:Contains: amino end of gag protein; env protein fragment; protein-tyrosine kinase
C:Species: avian leukosis virus, ALV
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-Oct-2004
C:Accession: B00643; A00643
R:Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines, M
Cell 41. 719-726, 1985
A>Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and pro
A:Reference number: A00643; MUID:85228222; PMID:2988784
A:Accession: B00643
A:Molecule type: mRNA
A:Residues: 1-698 <NID>
A:Cross-references: UNIPROT:P00534; UNIPARC:UPI000017113; GB:M10066; GB:M13881; NID:921
A>Note: in Genbank entry CHERBBF, release 109.0, the source is designated as Gallus gal
C:Comment: This protein is synthesized as a gag-env-erbB protein.
C:Genetic8:
A:Gene: gag-env-erbB
C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology
C:Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific P
F:1.6/Product: gag protein (fragment) #status predicted <DB>
F:7.59/Product: env protein (fragment) #status predicted <ENV>
F:60-698/Product: protein-tyrosine kinase erbB #status predicted <ERB>
F:194-459/Domain: protein-tyrosine kinase homology <KIN>
F:202-210/Region: protein kinase ATP-binding motif
F:229/Active site: Lys #status predicted
Query Match      25.9%; Score 1765.5; DB 1; Length 698;
Best Local Similarity 52.2%; Pred. No. 7.7e-67;
Matches 374; Conservative 80; Mismatches 137; Indels 125; Gaps 18;
578 G P E A D Q C V A C A H Y K D P P C V A R C P S G V K P D L S Y M P I M K F P E B G A C O P C P I N C T H S C V D L      637

```

```

Db      60  G P - - D H C M K A H F I D P H C V A C P A G V L G E N D T L - W K Y A D A N A V C O L C H R N C T R G C G P      116
Qy      638  D D K C P A E G R A S P L S I I S A V - G I L L V V T L G V V F G L I K R O O K I R Y T M R L L O E T E L      696
Db      117  G L E G C P - - - N G S K T P S I A G V A G G L C L V N G L G I G L T R R R - H I V K R T I R R L L Q E R L      172
Qy      697  V E P L T P S G A M P N O A M R I L K E T E L R K V L G S G A F G Y Y K K I W I P D G E N V K I P V A I K Y L R      756
Db      173  V E P L T P S G A M P N O A M R I L K E T E R K K Y L G S A F G Y Y K K I W I P E G K V K I P V A I K E L R      232
Qy      757  E N T S P K A N K E I L D A Y A V M A G V S P Y S R L G I C L T S T V Q L V T Q L M P R Y G C L L D H R E N G R      816
Db      233  E A T S P K A N K E I L D A Y A V M A S V D N P H V C R L G I C L T S T V Q L T Q L M P R Y G C L L D Y R E H K D N      292
Qy      817  L G S O D L M W C M Q I A K G S Y L E D V A L V H R D L A A R N V L V K S P N H V K I T D F G L A K L I G A D E K E      876
Db      293  I G S O Y L L M W C V O I A K G M Y L E E R L V H R D L A A R N V L V K T P Q H V K I T D F G L A K L I G A D E K E      352
Qy      877  Y H A D G K V P I K M A L E S I L R R F T H O S D V M S Y G V T V M E L M T F G A K P Y D G I P A R E I P D L L E      936
Db      353  Y H A E G K V P I K M A L E S I L R I Y T H O S D V M S Y G V T V M E L M T F G S K P Y D G I P A S E I S S V L E      412
Qy      937  K G E R L P O P I C T I D V M I M V K C M I D S E C R P R F E L V S F S R M A R D P O R F V I Q - N E D L G      995
Db      413  K G E R L P O P I C T I D V M I M V K C M I D A D S R K F R E L I A F S K M A R D P R Y L V I O G D E R M H      472
Qy      996  P A S P L D S T F Y S L I E D D M G L V D A E R Y L V Q O G F P C C D P A R A G A G M V H N H R S S T S G      1055
Db      473  L P S P L D S F Y T L M E E B E M E D I V D A E Y L V H Q G F - - - - - N S P S T - - -      513
Qy      1056  G G D L T L G L P E E E A P S P L - - - - - A P E D A G S V F P D G L G M A K G L S L P T H D P S P L Q      1110
Db      514  - - - - - S R P L L S L S A T S N N A T N C I D - - - - - - - - - - - - - - - - - - - - - - - - - - - -      550
Qy      1111  R Y S E D P T V L P S E T - D E Y A V L T C S P O R Y V N Q P D V A R P O P S P R E G P L A R A P A G A T L E      1168
Db      551  R Y S S D P T G N F L E S I D D G F L - - - - - P A B Y V N Q - - L M P K K R S - - - - - - - - - -      585
Qy      1169  R P K T L S P G K N G V A D V F - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -      1213
Db      586  - - - - - T A M Q N O I Y N N I S L T A I S K L P M D S R Y O N S H T A V D N E Y L - - - - - N T N Q S P L A      633
Qy      1214  S P A F D N L Y Y M D - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -      1254
Db      634  K T V F E S S F Y W I Q S G N H Q I N D N P O Y Q D F L P N E T K P N G L K V P A A E N P E Y L R V A P      689

RESULT 12
TVYU
protein-tyrosine kinase (EC 2.7.1.112) erbB - avian erythroblastosis virus (strain H)
C:Species: avian erythroblastosis virus
C>Date: 18-Apr-1984 #sequence_revision 18-Apr-1984 #text_change 05-Oct-2004
C:Accession: A00644; A38022
R:Yamamoto, T.; Nishida, T.; Miyajima, N.; Kawai, S.; Ooi, T.; Toyoshima, K.
Cell 35. 71-78, 1983
A>Title: The erbB gene of avian erythroblastosis virus is a member of the src gene famil
A:Reference number: A00644; MUID:84026539; PMID:6313229
A:Accession: A00644
A:Molecule type: DNA
A:Residues: 1-604 <YMA>
A:Cross-references: UNIPROT:P00535; UNIPARC:UPI000028A64; GB:K01216; NID:G209676; PIDN:
Science 224. 1456-1459, 1984
A>Title: Sequencing the erbA gene of avian erythroblastosis virus reveals a new type of
A:Reference number: A38022; MUID:84223957; PMID:6328658
A:Accession: A38022
A:Molecule type: DNA
A:Residues: 1-28, 'W', 30-139, 'F', 141-145, 'V', 147-152 <DBB>
A:Cross-references: UNIPARC:UPI00001725C9; GB:K02006
C:Genetic8:
A:Gene: erbB
C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology

```


C:Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific
F:130-395/Domain: protein kinase homology <KIN>
F:138-146/Region: protein kinase ATP-binding motif
F:165/Active site: Lys #status predicted

Query Match 25.0%; Score 1702; DB 1; Length 604;

Best Local Similarity 52.2%; Pred. No. 3e-64;

Matches 360; Conservative 76; Mismatches 128; Indels 126; Gaps 16;

QY 587 CAHKDPFCVACRCPGKPLSLVMPKRPDEGACQPCPINCCHSCVDDDDGCGPMEQ 646

Db 3 CAHFDGHCYKACAGVLENDTL-VAKYADANAVALCPKCTCKGKGLGCP--- 58

QY 647 RASPLTSLISAVV-GILLVVLGVVFGILLRQOKIRKRYTRRLDQETLVEPLPSGA 705

Db 59 NGSTPSTJAGVVGGLCLVVGIGIGLYARR-HIVAKRLRLRLQRELVPLPSGE 117

QY 706 MPNOQMRILKETELRKVKVLSGAFGTIVYGIWIPDGENYKIPVALVLRBENTSPKANK 765

Db 118 APNOAHRLIKETEKVKVLSGAFGTIVYGIWIPDGENYKIPVALVLRBENTSPKANK 177

QY 766 EILDEAYVMAVGSPPYVSRLLGLCTSTVOLVTOIMPGCLLDHYRNRRLGSGODLLNW 825

Db 178 EILDEAYVMAVGSPPYVSRLLGLCTSTVOLVTOIMPGCLLDHYRNRRLGSGODLLNW 237

QY 826 CMOJAKGMSYLEDEVRLVHRDLAARVVLVKSFNHYKITDFGLARLLDIDETRYHADGGKVP 885

Db 238 CVOJAKGMSYLEDEVRLVHRDLAARVVLVKSFNHYKITDFGLARLLDIDETRYHADGGKVP 297

QY 886 IKMALISILRRRTTHOSDWSYGVTTWELMTFGAKPYDGIIPABEIPDLLEKGRLLQPP 945

Db 298 IKMALISILRRRTTHOSDWSYGVTTWELMTFGAKPYDGIIPABEIPDLLEKGRLLQPP 357

QY 946 ICTDVMIMVCMWIDSECRPRELVSERSMARDDQRFVVIQ-NEDLGAPSLDSTF 1004

Db 358 ICTDVMIMVCMWIDSECRPRELVSERSMARDDQRFVVIQ-NEDLGAPSLDSTF 417

QY 1005 YRSLLEDMDGLVDAEYLVLPQGFPCRPDAPAGVGVHRRHRSSTSRSGGSDLLTGL 1064

Db 418 YRSLLEDMDGLVDAEYLVLPQGFPCRPDAPAGVGVHRRHRSSTSRSGGSDLLTGL 449

QY 1065 PSEBAPRSPV-----APSEAGSDVFPDGLGMAKAGLSPLPHDSPLQRYSEDPV 1119

Db 450 PSEBAPRSPV-----APSEAGSDVFPDGLGMAKAGLSPLPHDSPLQRYSEDPV 495

QY 1120 LPSEFT--DGVYAPLTCSPQPEYVQDPVQPPSPBEGRLPAAPAGATTLERPTLSPGK 1177

Db 496 PLEBSIDDFL-----PAPBYVNO--LMPKKPSTAM----- 524

QY 1178 NGVAKDVAF-----GGAVENTEYLTPQGAAPQHPAPAFSPAFD 1218

Db 525 --VQNGIYVNFISLVAISKLPMDSRKQNGSHSTAVDNPEYL-----NTNGSLAKTIVFE 574

QY 1219 NLTYWDDQPPRGAAPSTFKGTPFAENPEY 1248

Db 575 SSPYMIQSGNHQ-----INLDNDY 594

RESULT 13

Q:epidermal growth factor receptor - fruit fly (Drosophila melanogaster)

N:Contains: protein-tyrosine kinase (BC 2.7.1.112) erbr

C:Species: Drosophila melanogaster

C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 05-Oct-2004

C:Accession: A00640; A38021

R:Litven, E.; Glazer, L.; Segal, D.; Schlessinger, J.; Shilo, B.Z.

Cell 40; 599-607, 1985

A:Title: The Drosophila EGF receptor gene homolog: conservation of both hormone binding

A:Reference number: A00640; MUID:85124611; PMID:2982499

A:Molecule type: DNA

A:Residues: 1-1330 <LIV>

R:Madsworth, S.C.; Vincent III, W.S.; Blodéau-Wentworth, D.
Nature 314; 178-180, 1985

A:Title: A Drosophila genomic sequence with homology to human epidermal growth factor rec

A:Reference number: A38021; MUID:85137938; PMID:2983222

A:Accession: A38021

A:Molecule type: DNA

A:Residues: 'A', 832-866, 'V', 866-943, 'QTPSLVK' <MAD>

A:Cross-references: UNIPARC:UPI000016BBD5; EMBL:X02293; NID:97922; PID:CAA26157.1; PID:...

C:Comments: This sequence is tentative because the introns have not been identified.

C:Genetics:

A:Gene: FlyBase:Bgfr

A:Cross-references: FlyBase:FBgn0003731

A:Map position: 2 57F

C:Superfamily: Tyrosine-protein kinase, EGF receptor type; protein kinase homology

C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho

F:1732/Domain: extracellular #status predicted <EXT>

F:733-764/Domain: transmembrane #status predicted <TM>

F:765-1330/Domain: intracellular #status predicted <INT>

F:808-1072/Domain: protein kinase homology <KIN>

F:816-824/Region: protein kinase ATP-binding motif

F:122,300,324,363,518,688,695,700/Binding site: carbohydrate (Asn) (covalent) #status pr

F:774/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

F:843/Active site: Lys #status predicted

F:1181/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predict

Query Match 24.3%; Score 1653.5; DB 1; Length 1330;

Best Local Similarity 30.0%; Pred. No. 6.9e-62;

Matches 414; Conservative 179; Mismatches 415; Indels 371; Gaps 39;

QY 80 VQGVVLIAHQVROVPLQRLRIYRGTLF-----EDNYALAVLDNGPPLNTTIVTQASP 134

Db 38 ITNYIVIGLRLIPCTLSYRLQIIRGRLTFLSVSEETKAFV-----TY 81

QY 135 GGLREIOLRSITELTKGVLIOENPOLCYODTILMKDIFPKNNQALTLIDTNRBACHP 194

Db 82 SKMTLEIPLRLVNLQVGFHNHNYLCHKRTIOMSVISNGDYYNYDTABERCPK 141

QY 195 CSPWCKSGRCWGSSEDDQSLTRTVACAGCA--RCKGRLPTDCHEQCAAGCTGPKHSDC 252

Db 142 CHSCSTHG-CWGBQPKKQCKFSKLTCSPOCAGRCGYCPKPRECHLCAAGCTPTQKDC 200

QY 253 LACIHFHNSGICELHCAALVTYNTDTPESMPNBPGRRTFFGASCYTAPRYTLSDVSC 312

Db 201 IACNFFDEAVSKEECPPMKRYNTTYVLTNPBGKAYAGATCYKECP-GHLLENDACV 259

QY 313 LVCPFLHQEATBEDGTORCEKSCPCARVCYGLGMEHLREYRAVTSANIOEFAQCKIFG 372

Db 260 RSCPDODMDKGE-----CYPCNGPCPKTCPTGVVLA-----AGNIDSFRCYVYDG 306

QY 373 SLAFLEPSFDG--DPAANTA-----PLQPEQLQVFTLEITGYLYISAMPDSLPLDSV 424

Db 307 NIRLIDDTFGGFQDVVANYTMGPRIYPLDPERREVFSTVKEITGYLIEGTHPQFRILSY 366

QY 425 FQNLQVIRGRILNKAVAL-SITLQGLISWGLSLRLSGLALIHNTTLCPHYTPWD 483

Db 367 FRNLLETTHGRQLWESMFAALAIYKSLYSLEMRNLKQISSGSVVLIQNRDLCYVSNIRWP 426

QY 484 QLFENPHQALLHTANRPEDEC----- 504

Db 427 ALQKEPQKXVWVENLRADLCCKPFTLLISVQNHIIINHIFALCREKNHLLGSVQRRL 486

QY 505 ----- 504

Db 487 GSWHGVPLYQLQFQWHLHRLMLYIQVINSTQXSNHQLTDACYSBVSPTLTIER 546

QY 505 -----VEEGLA-- 510

Db 547 ARVALQAGLAMELEQITTAASAKRHSKTLPAEGQVPRWVFLGVCASAPAGIAPLAGR 606

QY 511 -----CHQLCARGHCWGPPTQCVNCSQPLRGQCEVECEVTLQGLPREVY---NARHCLP 562

Db 607 AVCRKCHPLCELCITNYIHQVCSKTHYGRREGCTEC-----PADHTYDEQRRCFQ 660

Query Match 24.1% Score 1645; DB 2; Length 544;
Best Local Similarity 57.7%; Pred. No. 6, 4e-62;
Matches 335; Conservative 65; Mismatch 97; Indels 84; Gaps 13;

QY 578 GBEADQVCVAHAKYKDPFCVACRPSGVKPDLSYMPKIFPDEBACQCPINCHSICYDL 637

Db 1 GP--DHCMCAHFIDBPHCVKACPAVGLGENDTL-VKTIADANAVCQLCHENCTRGCKGP 57

QY 638 DKGCPAEORASPLTSIIISAVV-GILLVVLGVVFGLIKRRQOKIRKYTRRLLOETEL 696

Db 58 GLEGCP---NGSKTPSIIAGVVGGLCLLVVGGIGIYLRR-HIVKRTLRRLLOEBEL 113

QY 697 VEPILPSGAMPNOQMRILKETELRKVKYLGSGAFGVYKGIWLPDGENVKIPVAIKYLR 756

Db 114 VEPILPSGAPNOAHRIKETEPRKVKYLGFGFVGVGLWLPBGEKVTIPVAIKELR 173

QY 757 EKTSPANKELIDEAYVMAVGSPYVSRLLGICLTSTVQJVTQJLMPYGLLDHYENRGR 816

Db 174 EKTSPANKELIDEAYVMAVDNPHYCRLLGICLTSTVQJLVTQJLMPYGLLDYIREHNDN 233

QY 817 LGSODLLMCMQIJAKMSYLEDLVLRHDLAARNVLKSPNHVKITDPGLARLLDIDETE 876

Db 234 IGSQYLLMNCVOJAKGMNYLEERHMYRDLAARNVLKTPQHVKITDPGLAKOLGADKE 293

QY 877 YHADGKVIIRKMALESILRRFTHOSDWSYGTVELMTFGAKPDGIDAREIPDLE 936

Db 294 YHAEQKVIIRKMALESILRRFTHOSDWSYGTVELMTFGSKPDGIDAREISSVLE 353

QY 937 KGERLPQPICTIDVYMIWKCMMIDSECRPRFRELSEFSRMRDPORFVYIQ-NEDLG 995

Db 354 KGERLPQPICTIDVYMIWKCMMDSADSRPKRELAEFSKMRDPPRYLVIOGDERMH 413

QY 996 PASPLDSTFYRSLLEDNDGDLVDAEYLVIPQOGFCPPDPAPGAGVHHRRSSSTRSG 1055

Db 414 LPSPTDSKYRTLMEBEDMEDIVDADEYLVPHQGF-----NSPST--- 454

QY 1056 GGDLTGLGEPSEEARSP-----APSEBAGSDVPFGDGMGAAGLQSLPTHPSPLO 1110

Db 455 -----SRTPLSLSLATSNSNATNCIDRNG-----H----- 481

QY 1111 RYSEDPVLPSTDGYVAPLTCSPOPEYVNOQDPVRQPPS 1151

Db 482 -----PVREDGFL-----PAPEYVNO-LMPKKRS 504

RESULT 15

S00727

kinase-related transforming protein (erbB) (EC 2.7.1.-) - avian erythroblastosis virus (

C:Species: avian erythroblastosis virus

C:Date: 01-Dec-1989 #sequence_rev=1989 01-Dec-1989 #text_change 31-Dec-2004

R:Accession: S00727

R:Scotting, P.; Vennstrom, B.; Jansen, M.; Graf, T.; Beug, H.; Hayman, M.J.

Oncogene Rev. 1, 265-278, 1987

A:Title: Common site of mutation in the erbB gene of avian erythroblastosis virus mutant

A:Reference number: S00727; MUID:88217326; PMID:2897102

A:Accession: S00727

A:Molecule type: DNA

A:Residues: 1-545 <SCOs>

A:Cross-references: UNIPROT:Q85468; UNIPROT:Q64895; UNIPARC:UPI000017561E; EMBL:X06943

C:Gene: erbB

C:Superfamily: protein kinase homology

C:Keywords: ATP; phosphotransferase

F:135-400/Domain: protein kinase homology <kin>

F:143-151/Region: protein kinase ATP-binding motif

Query Match 24.0%; Score 1638; DB 2; Length 545;

Best Local Similarity 57.7%; Pred. No. 1,3e-61;

Matches 335; Conservative 64; Mismatches 98; Indels 84; Gaps 13;

QY 578 GBEADQVCVAHAKYKDPFCVACRPSGVKPDLSYMPKIFPDEBACQCPINCHSICYDL 637

Db 1 GP--DHCMCAHFIDBPHCVKACPAVGLGENDTL-VKTIADANAVCQLCHENCTRGCKGP 57

QY 638 DKGCPAEORASPLTSIIISAVV-GILLVVLGVVFGLIKRRQOKIRKYTRRLLOETEL 696

Db 58 GLEGCP---NGSKTPSIIAGVVGGLCLLVVGGIGIYLRR-HIVKRTLRRLLOEBEL 113

QY 697 VEPILPSGAMPNOQMRILKETELRKVKYLGSGAFGVYKGIWLPDGENVKIPVAIKYLR 756

Db 114 VEPILPSGAPNOAHRIKETEPRKVKYLGFGFVGVGLWLPBGEKVTIPVAIKELR 173

QY 757 EKTSPANKELIDEAYVMAVGSPYVSRLLGICLTSTVQJVTQJLMPYGLLDHYENRGR 816

Db 174 EKTSPANKELIDEAYVMAVDNPHYCRLLGICLTSTVQJLVTQJLMPYGLLDYIREHNDN 233

QY 817 LGSODLLMCMQIJAKMSYLEDLVLRHDLAARNVLKSPNHVKITDPGLARLLDIDETE 876

Db 234 IGSQYLLMNCVOJAKGMNYLEERHMYRDLAARNVLKTPQHVKITDPGLAKOLGADKE 293

QY 877 YHADGKVIIRKMALESILRRFTHOSDWSYGTVELMTFGAKPDGIDAREIPDLE 936

Db 294 YHAEQKVIIRKMALESILRRFTHOSDWSYGTVELMTFGSKPDGIDAREISSVLE 353

QY 937 KGERLPQPICTIDVYMIWKCMMIDSECRPRFRELSEFSRMRDPORFVYIQ-NEDLG 995

Db 354 KGERLPQPICTIDVYMIWKCMMDSADSRPKRELAEFSKMRDPPRYLVIOGDERMH 413

QY 996 PASPLDSTFYRSLLEDNDGDLVDAEYLVIPQOGFCPPDPAPGAGVHHRRSSSTRSG 1055

Db 414 LPSPTDSKYRTLMEBEDMEDIVDADEYLVPHQGF-----NSPST--- 454

QY 1056 GGDLTGLGEPSEEARSP-----APSEBAGSDVPFGDGMGAAGLQSLPTHPSPLO 1110

Db 455 -----SRTPLSLSLATSNSNATNCIDRNG-----H----- 481

QY 1111 RYSEDPVLPSTDGYVAPLTCSPOPEYVNOQDPVRQPPS 1151

Db 482 -----PVREDGFL-----PAPEYVNO-LMPKKRS 504

638 DKGCPAEORASPLTSIIISAVV-GILLVVLGVVFGLIKRRQOKIRKYTRRLLOETEL 696

```
Db      58 GLEBGP---NGSKTSPSIAGVVGGLCLVVGGLGIGLYLRRR-HIVRRKRLRLLOEREL 113
Oy      697 VEPLTPSGAMPNOAOMRIKETELRKVKYLGSGAFGYKGIWIPDGENVKIPVAIKYLR 756
Db      114 VEPLTPSGEARNQAHRIKETEFKVKVYLGSGAFGYKGIWIPDGENVKIPVAIKELR 173
Oy      757 ENTSPKANKELIDEAYVNAVGSPPYSRLLGICLTSTVQVLTQIMPYGCLLDHVENRGR 816
Db      174 EATSPKANKELIDEAYVNAVSDNPHVCRLGICLTSTVQVLTQIMPYGCCLDYIREHKDN 233
Oy      817 LGSQDLLMWCQIAGMSYLEBVLVHRDLAARNVLVKSPPNHVKTITDGLARLDIDETE 876
Db      234 IGSQYLLMWCQIAGMNYLBERHLVHRDLAARNVLVKTPODVKITDGLAKOLGADEKE 293
Oy      877 YHADGKVPIKMMALESTILRRRFTQSDVMSYGVTVWELMTFGAKPYDGIPIAREIPDLLE 936
Db      294 YHABGKVPIKMMALESTILHRIYTHQSDVMSYGVTVWELMTFGSKPYDGIPIASEISSYLE 353
Oy      937 KGERLPOPPICTIDVYIMVCKMIDSECRPAFRELVSEFSRMARDPQRFVVIQ-NEDLG 995
Db      354 KGERLPOPPICTIDVYIMVCKMIDSECRPAFRELVSEFSRMARDPQRFVVIQ-NEDLG 995
Oy      996 PASPLDSTFYRSLEDDMGDLVDAEYLVPOQGFPCPDPAAGAGMWHHRHSSTRSG 1055
Db      414 LPSPTDSKFRYRLMEEDMEDIVDAEYLVPOQGFPCPDPAAGAGMWHHRHSSTRSG 1055
Oy      1056 GGDLLTGLEPSEBEAPRSPPL---APSEGASDVFDGDLGWAAGLQSLPTHDPSPLO 1110
Db      455 -----SRTPLLSSLSATSNNSATNCIDRNGC-----H----- 481
Oy      1111 RYSEBDPTVLPSETDGYVAPLTCSPQPEYVNOQDVVPQPPS 1151
Db      482 -----PVREDGFL-----PAPRYVNO-LMPKKPS 504
```

Search completed: January 3, 2006, 11:17:40
Job time : 59 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: January 3, 2006, 11:17:49 ; Search time 183 Seconds

(without alignments)
2865.442 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815

Sequence: 1 MELALCRWGLLLALLPPGA.....TFKGTPTAENPEYLGLDVPV 1255

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.dep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.dep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.dep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.dep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.dep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.dep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6815	100.0	1255	3	US-09-765-973-2
2	6815	100.0	1255	3	US-09-854-356-1
3	6815	100.0	1255	3	US-09-930-125-2
4	6815	100.0	1255	3	US-09-441-411-6
5	6815	100.0	1255	4	US-10-207-655-45
6	6815	100.0	1255	4	US-10-313-644-2
7	6815	100.0	1255	4	US-10-418-027-3
8	6815	100.0	1255	4	US-10-394-322A-17
9	6815	100.0	1255	4	US-10-245-871-553
10	6815	100.0	1255	4	US-10-469-162-3
11	6815	100.0	1255	4	US-10-253-286-553
12	6815	100.0	1255	5	US-10-762-128-6
13	6815	100.0	1255	5	US-10-484-067-1
14	6815	100.0	1255	5	US-10-723-860-9
15	6815	100.0	1255	5	US-10-871-708-9
16	6815	100.0	1255	5	US-10-983-340-12
17	6815	100.0	1255	6	US-11-037-713-12
18	6806	99.9	1255	3	US-09-811-123-9
19	6806	99.9	1255	3	US-09-811-115-3
20	6806	99.9	1255	3	US-09-354-533-68
21	6806	99.9	1255	4	US-09-984-092-4
22	6806	99.9	1255	4	US-10-177-293-126
23	6806	99.9	1255	4	US-10-207-498-6
24	6806	99.9	1255	4	US-10-338-730-2
25	6806	99.9	1255	4	US-10-322-892-4
26	6806	99.9	1255	4	US-10-272-437A-28
27	6806	99.9	1255	4	US-10-117-937-594

28	6806	99.9	1255	4	US-10-435-696-36	Sequence 36, Appl
29	6806	99.9	1255	4	US-10-647-005-68	Sequence 68, Appl
30	6806	99.9	1255	4	US-10-441-779C-4	Sequence 4, Appl
31	6806	99.9	1255	4	US-10-734-564-126	Sequence 126, App
32	6806	99.9	1255	4	US-10-657-022-90	Sequence 90, Appl
33	6806	99.9	1255	5	US-10-615-343-17	Sequence 17, Appl
34	6806	99.9	1255	5	US-10-794-514A-1	Sequence 1, Appl
35	6806	99.9	1255	5	US-10-846-113A-27	Sequence 27, Appl
36	6806	99.9	1255	5	US-11-067-064-594	Sequence 594, App
37	6806	99.9	1255	6	US-11-121-347-68	Sequence 68, Appl
38	6806	99.9	1255	6	US-11-067-159-594	Sequence 594, App
39	6804	99.8	1253	4	US-10-146-473-72	Sequence 72, Appl
40	6803	99.8	1255	4	US-10-149-138-4641	Sequence 4641, Ap
41	6803	99.8	1255	4	US-10-149-138-4641	Sequence 4641, Ap
42	6776.5	99.4	1294	5	US-10-956-373-10	Sequence 10, Appl
43	6757	99.1	1255	3	US-09-765-508-2	Sequence 2, Appl
44	6003	88.1	1257	5	US-10-484-067-2	Sequence 2, Appl
45	6001.5	88.1	1256	3	US-09-854-356-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1									
US-09-765-973-2									
; Sequence 2, Application US/09765973									
; Publication No. US20020039573A1									
; GENERAL INFORMATION:									
; APPLICANT: Cheever, Martin A.									
; TITLE OF INVENTION: Hand-Zimmerman, Susan									
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR PREVENTION AND									
; FILE REFERENCE: 210121.496									
; CURRENT APPLICATION NUMBER: US/09/765,973									
; CURRENT FILING DATE: 2001-01-19									
; NUMBER OF SEQ ID NOS: 4									
; SOFTWARE: FastSeq for Windows Version 3.0									
; SEQ ID NO 2									
; LENGTH: 1255									
; TYPR: PRT									
; ORGANISM: Homo sapien									
US-09-765-973-2									
Query Match									
100.0%; Score 6815; DB 3; Length 1255;									
Identical Similarity 100.0%; Pred. No. 0;									
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MELALCRWGLLLALLPPGA	STVCTGTDKRLPAS	PETHLDMRLHYGCGVVGNTL	60				
DB	1	MELALCRWGLLLALLPPGA	STVCTGTDKRLPAS	PETHLDMRLHYGCGVVGNTL	60				
QY	61	ELTYPTNASSFLDDIOEV	GVYLIANNQVQVPLQRL	RYRGVQLFEDNYALAVLNG	120				
DB	61	ELTYPTNASSFLDDIOEV	GVYLIANNQVQVPLQRL	RYRGVQLFEDNYALAVLNG	120				
QY	121	DLPLNTTVTASPGGLREL	QRLSTLTKGVLQRPOL	CYOPTIMKQIFHKNNOLA	180				
DB	121	DLPLNTTVTASPGGLREL	QRLSTLTKGVLQRPOL	CYOPTIMKQIFHKNNOLA	180				
QY	181	LTLIDTNRSRACHPCSP	CKSRGCRWESSBQOSL	TRTVACGACRCKGRLPTDC	240				
DB	181	LTLIDTNRSRACHPCSP	CKSRGCRWESSBQOSL	TRTVACGACRCKGRLPTDC	240				
QY	241	AAAGCTGPRGSDCLAL	HNHSGICELHCPALV	YNTDTFESMPNBEGRYTGAS	300				
DB	241	AAAGCTGPRGSDCLAL	HNHSGICELHCPALV	YNTDTFESMPNBEGRYTGAS	300				
QY	301	YNYLSTDVGSCTLVCP	LNQEVTAEDGTORCEK	SKPCARVCYGLGMEHLREVR	360				
DB	301	YNYLSTDVGSCTLVCP	LNQEVTAEDGTORCEK	SKPCARVCYGLGMEHLREVR	360				
QY	361	IOEFAGCKKIRGSLAF	LESFDGDBASTAPLQ	PEQLQVFTLBEITCYLI	420				
DB	361	IOEFAGCKKIRGSLAF	LESFDGDBASTAPLQ	PEQLQVFTLBEITCYLI	420				

Db 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLOPELOVFEETLEITGYLISAMPDLSL 420
Qy 421 DLSVFONLQVIRGRILHNGAYSLTLQGLGISLHGRSLRELSGSLALIHNTHLCEFYATV 480
Db 421 DLSVFONLQVIRGRILHNGAYSLTLQGLGISLHGRSLRELSGSLALIHNTHLCEFYATV 480
Qy 481 PWDOLFRRPHQALHTANRPEDECVGEGLAHQLCARHGWGPGPTQCVNCSQFLRGEC 540
Db 481 PWDOLFRRPHQALHTANRPEDECVGEGLAHQLCARHGWGPGPTQCVNCSQFLRGEC 540
Qy 541 VEECVVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCACAHYADPPFCVARC 600
Db 541 VEECVVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCACAHYADPPFCVARC 600
Qy 601 PSGVPPDLSYMPIMKFPDEBGACOPCPINCHTSCVDLDDKCPABORASPLTISIAYVG 660
Db 601 PSGVPPDLSYMPIMKFPDEBGACOPCPINCHTSCVDLDDKCPABORASPLTISIAYVG 660
Qy 661 ILVVVLOGVFGILLIKRQOKIRKXTMRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
Db 661 ILVVVLOGVFGILLIKRQOKIRKXTMRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
Qy 721 RKVVLGSGARGRTYKGIWIPRGENVKIPVAIKULRENTSPKANKEIIDELAYVAGVSP 780
Db 721 RKVVLGSGARGRTYKGIWIPRGENVKIPVAIKULRENTSPKANKEIIDELAYVAGVSP 780
Qy 781 YVSRLLGICLTSTVOLTQMLPVGCLLDHVRNRRGLSODLLNMCQIAKMSYLEDVR 840
Db 781 YVSRLLGICLTSTVOLTQMLPVGCLLDHVRNRRGLSODLLNMCQIAKMSYLEDVR 840
Qy 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETRYHADGGKVPDKMMALLESILRRFT 900
Db 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETRYHADGGKVPDKMMALLESILRRFT 900
Qy 901 HOSDWMSGYVWELMTGAKRYDGIIPAREIPDLLEKERRLPQRPICITDIYVMIVKWM 960
Db 901 HOSDWMSGYVWELMTGAKRYDGIIPAREIPDLLEKERRLPQRPICITDIYVMIVKWM 960
Qy 961 IDSECRPFRELVSFSSMARDPQRFVYIQNEDLPASPLDSTFYRSLLEDMDKDLVDA 1020
Db 961 IDSECRPFRELVSFSSMARDPQRFVYIQNEDLPASPLDSTFYRSLLEDMDKDLVDA 1020
Qy 1021 EBYLVPOQGFPCPDPAFGAGVNHRRHSSSTRSGGDLITGLPESSEBAPRSLAPSEG 1080
Db 1021 EBYLVPOQGFPCPDPAFGAGVNHRRHSSSTRSGGDLITGLPESSEBAPRSLAPSEG 1080
Qy 1081 AGSDVFPDDDLGMAKGLQSLPTHPSPLOQRYSEDPVPLBSETGVAVPLTCSQPPEYV 1140
Db 1081 AGSDVFPDDDLGMAKGLQSLPTHPSPLOQRYSEDPVPLBSETGVAVPLTCSQPPEYV 1140
Qy 1141 NOPDVRPOPSPREBGLPAPARPAAGATLERPKTLSEGKGVVADVAFGAVENPEYLTPO 1200
Db 1141 NOPDVRPOPSPREBGLPAPARPAAGATLERPKTLSEGKGVVADVAFGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPAFDNLYYWDODPBERGAPBSTFKGTPTAENPEYLGIDVPV 1255
Db 1201 GGAAPQHPHPAPAFDNLYYWDODPBERGAPBSTFKGTPTAENPEYLGIDVPV 1255

RESULT 2
US-09-854-356-1

; Sequence 1, Application US/09854356
; Patent No. US020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheyben, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: Smithkline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; CURRENT FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480

; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1255
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human HER-2/neu protein
; NAME/KEY: DOMAIN
; LOCATION: (1)..(653)
; OTHER INFORMATION: extracellular domain (ECD)
; NAME/KEY: DOMAIN
; LOCATION: (676)..(1255)
; OTHER INFORMATION: intracellular domain (ICD)
; NAME/KEY: DOMAIN
; LOCATION: (990)..(1255)
; OTHER INFORMATION: phosphorylation domain (PD)
; NAME/KEY: DOMAIN
; LOCATION: (990)..(1048)
; OTHER INFORMATION: fragment of the phosphorylation domain, preferred
; OTHER INFORMATION: portion (delta PD)
US-09-854-356-1

Query Match 100.0%; Score 6815; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MELALCRWGLLLALPPGAASSTOVCTGDMKRLPASPEYTHLMLRHLVGGCVVQGNL 60
Db 1 MELALCRWGLLLALPPGAASSTOVCTGDMKRLPASPEYTHLMLRHLVGGCVVQGNL 60
Qy 61 ELTYLPNASISFLQDIQEVGVYLIHNOVROVPLRLIVRGTOLEPDNYALAVDNG 120
Db 61 ELTYLPNASISFLQDIQEVGVYLIHNOVROVPLRLIVRGTOLEPDNYALAVDNG 120
Qy 121 DPLNNTTPVVGASPGGLRELQRLSTELKGVLIQRPOLCYODTILMKDI FHKNNQLA 180
Db 121 DPLNNTTPVVGASPGGLRELQRLSTELKGVLIQRPOLCYODTILMKDI FHKNNQLA 180
Qy 181 LTLIDTNRSRACHPCSPMCKGSRGSSSEDCQSLTRTVCCGGCARCKGRLPTCCHEQC 240
Db 181 LTLIDTNRSRACHPCSPMCKGSRGSSSEDCQSLTRTVCCGGCARCKGRLPTCCHEQC 240
Qy 241 AAGCTGPKHSQCLACLPNHSIGICELHCPALVTYNTDFFESMPNBEGRYTGASCVTACP 300
Db 241 AAGCTGPKHSQCLACLPNHSIGICELHCPALVTYNTDFFESMPNBEGRYTGASCVTACP 300
Qy 301 YNYLSTDVGSCTLYVCPHNDQVTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
Db 301 YNYLSTDVGSCTLYVCPHNDQVTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
Qy 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLOPELOVFEETLEITGYLISAMPDLSL 420
Db 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLOPELOVFEETLEITGYLISAMPDLSL 420
Qy 421 DLSVFONLQVIRGRILHNGAYSLTLQGLGISLHGRSLRELSGSLALIHNTHLCEFYATV 480
Db 421 DLSVFONLQVIRGRILHNGAYSLTLQGLGISLHGRSLRELSGSLALIHNTHLCEFYATV 480
Qy 481 PWDOLFRRPHQALHTANRPEDECVGEGLAHQLCARHGWGPGPTQCVNCSQFLRGEC 540
Db 481 PWDOLFRRPHQALHTANRPEDECVGEGLAHQLCARHGWGPGPTQCVNCSQFLRGEC 540
Qy 541 VEECVVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCACAHYADPPFCVARC 600
Db 541 VEECVVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQCACAHYADPPFCVARC 600
Qy 601 PSGVPPDLSYMPIMKFPDEBGACOPCPINCHTSCVDLDDKCPABORASPLTISIAYVG 660
Db 601 PSGVPPDLSYMPIMKFPDEBGACOPCPINCHTSCVDLDDKCPABORASPLTISIAYVG 660

QY 661 ILVVVLGVVFGILLIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQRILKETEL 720
DB 661 ILVVVLGVVFGILLIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQRILKETEL 720
QY 721 RKVKVLSGAGFTYVKGIMI PDGENVKIPVAIKVIRENTSPKANKEILDEAYVMAVGSP 780
DB 721 RKVKVLSGAGFTYVKGIMI PDGENVKIPVAIKVIRENTSPKANKEILDEAYVMAVGSP 780
QY 781 YVSRLLGICLTSTVQLVQMLPYGCLLDHVENRGRGLSQDILNMCQIAKMSYLEDR 840
DB 781 YVSRLLGICLTSTVQLVQMLPYGCLLDHVENRGRGLSQDILNMCQIAKMSYLEDR 840
QY 841 LVHRDLAARNVLYKSPNHVKTIDFGLARLLIDETRYADGKVPKIMMALESILRRFT 900
DB 841 LVHRDLAARNVLYKSPNHVKTIDFGLARLLIDETRYADGKVPKIMMALESILRRFT 900
QY 901 HQSDVMSYGVTVWELMTFGAKPYDGI PARBIPDLLEKGERLPQPICTIDVYMIWCKM 960
DB 901 HQSDVMSYGVTVWELMTFGAKPYDGI PARBIPDLLEKGERLPQPICTIDVYMIWCKM 960
QY 961 IDSECRPRRELVSFSPMAADPQRFVVIQNEDLGPASPLDSTFYRSLDEDDMGDLVDA 1020
DB 961 IDSECRPRRELVSFSPMAADPQRFVVIQNEDLGPASPLDSTFYRSLDEDDMGDLVDA 1020
QY 1021 BEYLVPOQGFPCPDPAFGAGVHHRRSSSTRSGGDLTLGLPSEBEPASPLAPSEG 1080
DB 1021 BEYLVPOQGFPCPDPAFGAGVHHRRSSSTRSGGDLTLGLPSEBEPASPLAPSEG 1080
QY 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140
DB 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140
QY 1141 NOPDVRPQPSREGBPLPAAPAGATLERPKTSLSGKGVKQVAVAFGAVENPEYLTPQ 1200
DB 1141 NOPDVRPQPSREGBPLPAAPAGATLERPKTSLSGKGVKQVAVAFGAVENPEYLTPQ 1200
QY 1201 GGAAPQHPPPAFSPAFDNLVYMDODPPERGA PPSFTFKGTPTAENPEYLGDLVPV 1255
DB 1201 GGAAPQHPPPAFSPAFDNLVYMDODPPERGA PPSFTFKGTPTAENPEYLGDLVPV 1255

RESULT 3
US-09-930-125-2
Sequence 2, Application US/09930125
Publication No. US20020193329A1
GENERAL INFORMATION:
APPLICANT: Hand-Zimmerman, Susan
APPLICANT: Cheever, Martin A.
APPLICANT: Foy, Teresa M.
APPLICANT: Lodge, Michael J.
APPLICANT: Kates, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: Vedvick, Thomas S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.544
CURRENT APPLICATION NUMBER: US/09/930,125
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapien
US-09-930-125-2

Query Match 100.0%; Score 6815; DB 3; Length 1255;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METALACRWGLLALLPFGAASVQCTGTDMKRLRPASBETHLDMRLHLYGCGQVVGNTL 60
DB 1 METALACRWGLLALLPFGAASVQCTGTDMKRLRPASBETHLDMRLHLYGCGQVVGNTL 60

DB 1 METALACRWGLLALLPFGAASVQCTGTDMKRLRPASBETHLDMRLHLYGCGQVVGNTL 60
QY 61 ELTYLPTNASISFLDIOEVQGYVLIAHNVQVPELORLRIYRGTLFEDNYALAVLDNG 120
DB 61 ELTYLPTNASISFLDIOEVQGYVLIAHNVQVPELORLRIYRGTLFEDNYALAVLDNG 120
QY 121 DPLANTTIVTASPGGLREBLQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNQDLA 180
DB 121 DPLANTTIVTASPGGLREBLQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNQDLA 180
QY 181 LTLIDTNRSRACHPCSPMCKSGRCWSESSDDQSILTRIVYACGGACRCGRLPTDCHBQC 240
DB 181 LTLIDTNRSRACHPCSPMCKSGRCWSESSDDQSILTRIVYACGGACRCGRLPTDCHBQC 240
QY 241 AAGCTGPHNSDCLAETHNHSICELHCPALVTYNTDFESMNPNEGRTFGASCVTACP 300
DB 241 AAGCTGPHNSDCLAETHNHSICELHCPALVTYNTDFESMNPNEGRTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPLNQEVTAEDGTORCEKSKPCARVYCYGMEHLREVRVTSAN 360
DB 301 YNYLSTDVGSCTLVCPLNQEVTAEDGTORCEKSKPCARVYCYGMEHLREVRVTSAN 360
QY 361 IOEFAGCKKIFGSLAFLPBSPDGPASNTAPLOEPQLOVFTLEBITGYLISAMPDSLIP 420
DB 361 IOEFAGCKKIFGSLAFLPBSPDGPASNTAPLOEPQLOVFTLEBITGYLISAMPDSLIP 420
QY 421 DLSVFQNTQVTRGRILHNAGASLTLOGIGISWLGRLSRLBSSGLALIHNNHILCFVHTV 480
DB 421 DLSVFQNTQVTRGRILHNAGASLTLOGIGISWLGRLSRLBSSGLALIHNNHILCFVHTV 480
QY 481 PWDQLFRPHQALLHTANRPEDECEVGEGLACHQLCARGHCMPGTQCVCNSQFLRGGEC 540
DB 481 PWDQLFRPHQALLHTANRPEDECEVGEGLACHQLCARGHCMPGTQCVCNSQFLRGGEC 540
QY 541 VEECKVLOGLPREYVNAHCLPCHBECOPQNGSVTCFPEADQCAVACHYKOPPCVAVRC 600
DB 541 VEECKVLOGLPREYVNAHCLPCHBECOPQNGSVTCFPEADQCAVACHYKOPPCVAVRC 600
QY 601 PSGVCPDLSYMPIMKFPBEBGACOPCTNGCHSCVDLDDKGPABORASPLTSIIISAVYG 660
DB 601 PSGVCPDLSYMPIMKFPBEBGACOPCTNGCHSCVDLDDKGPABORASPLTSIIISAVYG 660
QY 661 ILVVVLGVVFGILLIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQRILKETEL 720
DB 661 ILVVVLGVVFGILLIKRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQRILKETEL 720
QY 721 RKVKVLSGAGFTYVKGIMI PDGENVKIPVAIKVIRENTSPKANKEILDEAYVMAVGSP 780
DB 721 RKVKVLSGAGFTYVKGIMI PDGENVKIPVAIKVIRENTSPKANKEILDEAYVMAVGSP 780
QY 781 YVSRLLGICLTSTVQLVQMLPYGCLLDHVENRGRGLSQDILNMCQIAKMSYLEDR 840
DB 781 YVSRLLGICLTSTVQLVQMLPYGCLLDHVENRGRGLSQDILNMCQIAKMSYLEDR 840
QY 841 LVHRDLAARNVLYKSPNHVKTIDFGLARLLIDETRYADGKVPKIMMALESILRRFT 900
DB 841 LVHRDLAARNVLYKSPNHVKTIDFGLARLLIDETRYADGKVPKIMMALESILRRFT 900
QY 901 HQSDVMSYGVTVWELMTFGAKPYDGI PARBIPDLLEKGERLPQPICTIDVYMIWCKM 960
DB 901 HQSDVMSYGVTVWELMTFGAKPYDGI PARBIPDLLEKGERLPQPICTIDVYMIWCKM 960
QY 961 IDSECRPRRELVSFSPMAADPQRFVVIQNEDLGPASPLDSTFYRSLDEDDMGDLVDA 1020
DB 961 IDSECRPRRELVSFSPMAADPQRFVVIQNEDLGPASPLDSTFYRSLDEDDMGDLVDA 1020
QY 1021 BEYLVPOQGFPCPDPAFGAGVHHRRSSSTRSGGDLTLGLPSEBEPASPLAPSEG 1080
DB 1021 BEYLVPOQGFPCPDPAFGAGVHHRRSSSTRSGGDLTLGLPSEBEPASPLAPSEG 1080
QY 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140
DB 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140

```

Oy 1141 NOPVRRPQPSRREGPLPAARPAAGATLBRPKTSLSPGKGVVWDVAFGAVENBEYLTPQ 1200
Db 1141 NOPVRRPQPSRREGPLPAARPAAGATLBRPKTSLSPGKGVVWDVAFGAVENBEYLTPQ 1200
Oy 1201 GGAAPQHPPPAFSPAFDNLVYWDQDPPERGAPESTFKGTPTTANPEYLGIDVY 1255
Db 1201 GGAAPQHPPPAFSPAFDNLVYWDQDPPERGAPESTFKGTPTTANPEYLGIDVY 1255

RESULT 4
US-09-441-411-6
; Sequence 6, Application US/09441411
; Publication No. US2003008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disib, Mary L.
; APPLICANT: Helistrom, Ingegerd
; APPLICANT: Helistrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 73003.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-411-6

Query Match 100.0%; Score 6815; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MELAALCRWGILLALLPPGAASSTQVCTGDMKRLPASPEETHLDMRLHLYGSCQVQGNL 60
Db 1 MELAALCRWGILLALLPPGAASSTQVCTGDMKRLPASPEETHLDMRLHLYGSCQVQGNL 60
Oy 61 ELTYLPTNASLSFODIODEVQGVLIANNOVQVLOGLRIYRGTOQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFODIODEVQGVLIANNOVQVLOGLRIYRGTOQLFEDNYALAVLDNG 120
Oy 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQNRNPOLCYODTIIMKDI FHNQOLA 180
Db 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQNRNPOLCYODTIIMKDI FHNQOLA 180
Oy 181 LTLIDTNRSRACHPCSPMKSGRCSGESSEDCQSLTRTVCAAGGCARCKGPLEPTDCHEQC 240
Db 181 LTLIDTNRSRACHPCSPMKSGRCSGESSEDCQSLTRTVCAAGGCARCKGPLEPTDCHEQC 240
Oy 241 AAGCGPRHSDCLALHFHNSGICELHCPALVTYNTDFESHPNBEGRYTGASCVTACP 300
Db 241 AAGCGPRHSDCLALHFHNSGICELHCPALVTYNTDFESHPNBEGRYTGASCVTACP 300
Oy 301 YNYLSTDVGSCTLVPLHNOEVTAEADGTQRCCKSKPCARVCYGLGMEHLREVRATISAN 360
Db 301 YNYLSTDVGSCTLVPLHNOEVTAEADGTQRCCKSKPCARVCYGLGMEHLREVRATISAN 360
Oy 361 IOEFAGCKKIFGSLAFLEPESFDGDPASNTAPLQPOLOVFFETLEBITGYLISAMPDSL 420
Db 361 IOEFAGCKKIFGSLAFLEPESFDGDPASNTAPLQPOLOVFFETLEBITGYLISAMPDSL 420
Oy 421 DLSVQNTQVTRGRILHNGAVSLTQAGISWLGRSLRELGSGALIHNNTHLCFVHTV 480
Db 421 DLSVQNTQVTRGRILHNGAVSLTQAGISWLGRSLRELGSGALIHNNTHLCFVHTV 480
Oy 481 PMDOFLRPHQALHTANRPEDECVGEGLACHQLCARGHCWPGPTQCVNCSQFLRGQEC 540
Db 481 PMDOFLRPHQALHTANRPEDECVGEGLACHQLCARGHCWPGPTQCVNCSQFLRGQEC 540
Oy 541 VEECRVLOGLPREYVNAHRCLEPCECOPNGSVTCFGEADQVCAAHYKDPFCVARC 600
Db 541 VEECRVLOGLPREYVNAHRCLEPCECOPNGSVTCFGEADQVCAAHYKDPFCVARC 600
```

```

Db 541 VEECRVLOGLPREYVNAHRCLEPCECOPNGSVTCFGEADQVCAAHYKDPFCVARC 600
Oy 601 PSQVPLSLYMPILWKFPEDEGACQPCPINCTHSCVDLDDKCAPEORASPLTISAVVG 660
Db 601 PSQVPLSLYMPILWKFPEDEGACQPCPINCTHSCVDLDDKCAPEORASPLTISAVVG 660
Oy 661 ILLVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMRYOAMRLIKSTEL 720
Db 661 ILLVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMRYOAMRLIKSTEL 720
Oy 721 RKVVLGSGAFGYVKGIWI PDGENVKI PVAIKVLRNTSPKAKEILDEAYVAGVGP 780
Db 721 RKVVLGSGAFGYVKGIWI PDGENVKI PVAIKVLRNTSPKAKEILDEAYVAGVGP 780
Oy 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHYRENGRLGSDOLIMWCMQIAKMSYLEDDR 840
Db 781 YVSRLLGICLTSTVQVLTQMLPYGCLLDHYRENGRLGSDOLIMWCMQIAKMSYLEDDR 840
Oy 841 LVHRDLAARVNLVKSNNHVKITDPGLARLDDIDETBYHADGKVPIKMALESILRRFT 900
Db 841 LVHRDLAARVNLVKSNNHVKITDPGLARLDDIDETBYHADGKVPIKMALESILRRFT 900
Oy 901 HQSDVMSYGVTVMEIEMFGAKPVYDGIIPAREIPDLLEKGERLPQPPICITIDVYIMVCKM 960
Db 901 HQSDVMSYGVTVMEIEMFGAKPVYDGIIPAREIPDLLEKGERLPQPPICITIDVYIMVCKM 960
Oy 961 IDSECRPRFRELVESEFRMARADPQRFVYIQNEDIGPASPDLSTFYRSLDEDDMDGLVDA 1020
Db 961 IDSECRPRFRELVESEFRMARADPQRFVYIQNEDIGPASPDLSTFYRSLDEDDMDGLVDA 1020
Oy 1021 EBYLVPOQGFPCPDPAAGAGMHHRRSSSTRSGGDLTLGLEPSEBEARSLAASEG 1080
Db 1021 EBYLVPOQGFPCPDPAAGAGMHHRRSSSTRSGGDLTLGLEPSEBEARSLAASEG 1080
Oy 1081 AGSDVPDGDIGMGAKEGLQSLPTHDSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYV 1140
Db 1081 AGSDVPDGDIGMGAKEGLQSLPTHDSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYV 1140
Oy 1141 NOPVRRPQPSRREGPLPAARPAAGATLBRPKTSLPGKGVVWDVAFGAVENBEYLTPQ 1200
Db 1141 NOPVRRPQPSRREGPLPAARPAAGATLBRPKTSLPGKGVVWDVAFGAVENBEYLTPQ 1200
Oy 1201 GGAAPQHPPPAFSPAFDNLVYWDQDPPERGAPESTFKGTPTTANPEYLGIDVY 1255
Db 1201 GGAAPQHPPPAFSPAFDNLVYWDQDPPERGAPESTFKGTPTTANPEYLGIDVY 1255

RESULT 5
US-10-207-655-45
; Sequence 45, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-45

Query Match 100.0%; Score 6815; DB 4; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MELAALCRWGILLALLPPGAASSTQVCTGDMKRLPASPEETHLDMRLHLYGSCQVQGNL 60
Db 1 MELAALCRWGILLALLPPGAASSTQVCTGDMKRLPASPEETHLDMRLHLYGSCQVQGNL 60
```


QY 61 ELTYLPTNASISFLODIOEVGVYLIAHQVROVPLQRLIRYRGTOLEFEDNYALAVLDNG 120
 DB 61 ELTYLPTNASISFLODIOEVGVYLIAHQVROVPLQRLIRYRGTOLEFEDNYALAVLDNG 120
 QY 121 DPLANTTPTVGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNQOLA 180
 DB 121 DPLANTTPTVGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNQOLA 180
 QY 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCOSLTRTYCAGGCARCKPLPTDCCHQC 240
 DB 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCOSLTRTYCAGGCARCKPLPTDCCHQC 240
 QY 241 AAGCTGPRKSDCLAHFNHSGICEIHCPCALVTYNTDFESMPNREGRTTFGASCVTACP 300
 DB 241 AAGCTGPRKSDCLAHFNHSGICEIHCPCALVTYNTDFESMPNREGRTTFGASCVTACP 300
 QY 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
 DB 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
 QY 361 IOEPFAGCKKIFGSLAFLPESFDGDPASNTAPLOPQLOVFTLEETITGYLISAMPDSL 420
 DB 361 IOEPFAGCKKIFGSLAFLPESFDGDPASNTAPLOPQLOVFTLEETITGYLISAMPDSL 420
 QY 421 DLSVFQNLQVIRGRILHNGAYSILTQGLGISWLGIRSLRELDSGLALIHNTHLCFVHTV 480
 DB 421 DLSVFQNLQVIRGRILHNGAYSILTQGLGISWLGIRSLRELDSGLALIHNTHLCFVHTV 480
 QY 481 PWDQLFRNPQALHTTANRPEDECVGEGIALCHQLCARHGCWGPPTQCVNCSQFLRGQEC 540
 DB 481 PWDQLFRNPQALHTTANRPEDECVGEGIALCHQLCARHGCWGPPTQCVNCSQFLRGQEC 540
 QY 541 VEECVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCVACAHYKDPPECVARC 600
 DB 541 VEECVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCVACAHYKDPPECVARC 600
 QY 601 PSGVPRDLSYMPIMKFPDEBGACOPCLNCTHSQVDLDKGCAPAGORASPLSIISAVVG 660
 DB 601 PSGVPRDLSYMPIMKFPDEBGACOPCLNCTHSQVDLDKGCAPAGORASPLSIISAVVG 660
 QY 720 ILLVVLGVFGIILIKRQOKIRKTYTMRILQETELVEPLTPSGAMPNOQRILKETEL 720
 DB 720 ILLVVLGVFGIILIKRQOKIRKTYTMRILQETELVEPLTPSGAMPNOQRILKETEL 720
 QY 780 RKVLYLGSAGFTYKGIWIPDGENVKIPVAIKVIRENTSPYANKREILDEAYVMAGVGP 780
 DB 780 RKVLYLGSAGFTYKGIWIPDGENVKIPVAIKVIRENTSPYANKREILDEAYVMAGVGP 780
 QY 840 YVSRLLGICLTSTVOLVQMLPYGCLLDHVRNRRGLSQDILNMCQIAKMSYLEDR 840
 DB 840 YVSRLLGICLTSTVOLVQMLPYGCLLDHVRNRRGLSQDILNMCQIAKMSYLEDR 840
 QY 900 LVHRDLAARNVLYKSPNHVKITDQGLARLLDIDETEHADGKVPDKMMALLESILRRRT 900
 DB 900 LVHRDLAARNVLYKSPNHVKITDQGLARLLDIDETEHADGKVPDKMMALLESILRRRT 900
 QY 960 HOSDWMSYGVTVWEIMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVTMINKCM 960
 DB 960 HOSDWMSYGVTVWEIMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVTMINKCM 960
 QY 1020 IDSECRPRRELVSERSMARDPORFVVIQNEIDLGPASPLDSTFRLSLEDMDGLVDA 1020
 DB 1020 IDSECRPRRELVSERSMARDPORFVVIQNEIDLGPASPLDSTFRLSLEDMDGLVDA 1020
 QY 1080 BEYLVPQOGFCPCPDPAAGAGVHHRRSSSTRSGGDLTLGLBESSEBAPRSLAPSEG 1080
 DB 1080 BEYLVPQOGFCPCPDPAAGAGVHHRRSSSTRSGGDLTLGLBESSEBAPRSLAPSEG 1080
 QY 1140 AGSDVFDGDLGMAKAGLOSLPTHDPSPLOKYSBDPTVPLPSETDGYVALPLTCSPOPEV 1140
 DB 1140 AGSDVFDGDLGMAKAGLOSLPTHDPSPLOKYSBDPTVPLPSETDGYVALPLTCSPOPEV 1140

QY 1141 NOPDVRPOPSPBREGPLPAAPAGATLIERPKTLSPGKNQVNDVAFGGAIVENPEYLTPQ 1200
 DB 1141 NOPDVRPOPSPBREGPLPAAPAGATLIERPKTLSPGKNQVNDVAFGGAIVENPEYLTPQ 1200
 QY 1201 GGAAPQHPPPAFSAFONLYTWDODPREBGPSTFTFGTPTAENPEYLGIDVP 1255
 DB 1201 GGAAPQHPPPAFSAFONLYTWDODPREBGPSTFTFGTPTAENPEYLGIDVP 1255
 RESULT 6
 US-10-313-644-2
 ; Sequence 2, Application US/10313644
 ; Publication No. US20030157119A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Cheever, Martin A.
 ; APPLICANT: Hand-Zimmerman, Susan
 ; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF HEMATOLOGICAL
 ; FILE REFERENCE: 210121.483C3
 ; CURRENT APPLICATION NUMBER: US/10/313,644
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 1255
 ; TYPE: PRF
 ; ORGANISM: Homo sapien
 ; US-10-313-644-2

Query Match 100.0%; Score 6815; DB 4; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELALCWGGLIALLPFGAASVOYCTGTDMKLRIPASBETHLDMRLHYGCCOYVQGNL 60
 DB 1 MELALCWGGLIALLPFGAASVOYCTGTDMKLRIPASBETHLDMRLHYGCCOYVQGNL 60
 QY 61 ELTYLPTNASISFLODIOEVGVYLIAHQVROVPLQRLIRYRGTOLEFEDNYALAVLDNG 120
 DB 61 ELTYLPTNASISFLODIOEVGVYLIAHQVROVPLQRLIRYRGTOLEFEDNYALAVLDNG 120
 QY 121 DPLANTTPTVGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNQOLA 180
 DB 121 DPLANTTPTVGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNQOLA 180
 QY 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCOSLTRTYCAGGCARCKPLPTDCCHQC 240
 DB 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCOSLTRTYCAGGCARCKPLPTDCCHQC 240
 QY 241 AAGCTGPRKSDCLAHFNHSGICEIHCPCALVTYNTDFESMPNREGRTTFGASCVTACP 300
 DB 241 AAGCTGPRKSDCLAHFNHSGICEIHCPCALVTYNTDFESMPNREGRTTFGASCVTACP 300
 QY 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
 DB 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
 QY 361 IOEPFAGCKKIFGSLAFLPESFDGDPASNTAPLOPQLOVFTLEETITGYLISAMPDSL 420
 DB 361 IOEPFAGCKKIFGSLAFLPESFDGDPASNTAPLOPQLOVFTLEETITGYLISAMPDSL 420
 QY 421 DLSVFQNLQVIRGRILHNGAYSILTQGLGISWLGIRSLRELDSGLALIHNTHLCFVHTV 480
 DB 421 DLSVFQNLQVIRGRILHNGAYSILTQGLGISWLGIRSLRELDSGLALIHNTHLCFVHTV 480
 QY 481 PWDQLFRNPQALHTTANRPEDECVGEGIALCHQLCARHGCWGPPTQCVNCSQFLRGQEC 540
 DB 481 PWDQLFRNPQALHTTANRPEDECVGEGIALCHQLCARHGCWGPPTQCVNCSQFLRGQEC 540
 QY 541 VEECVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCVACAHYKDPPECVARC 600
 DB 541 VEECVLQGLPREYVNAHCLPCHPECCOPONGSVTCFGEADQCVACAHYKDPPECVARC 600

Qy	601	PSGVKPDLSYMPIMKFPPEBGAOCPCPINCHTSCVDLDDKCGPAORASPLTSTISAVYG	660
Db	601	PSGVKPDLSYMPIMKFPPEBGAOCPCPINCHTSCVDLDDKCGPAORASPLTSTISAVYG	660
Qy	661	ILVVVLGVGVGIIILKRRQOKIRKTYMRLLOETLVEPLFPSGAMPNOQWRILKETEL	720
Db	661	ILVVVLGVGVGIIILKRRQOKIRKTYMRLLOETLVEPLFPSGAMPNOQWRILKETEL	720
Qy	721	RKVVLSSGAGFCTVYKGIWIPDGENVKIPVAIKVIRENTSPRANKIIDEAVYMGVSP	780
Db	721	RKVVLSSGAGFCTVYKGIWIPDGENVKIPVAIKVIRENTSPRANKIIDEAVYMGVSP	780
Qy	781	YVSRLLGICLTSTVOLVQMLMPYGLLDHVRENRRGLSODLLNMCQIAGKMSYLEVR	840
Db	781	YVSRLLGICLTSTVOLVQMLMPYGLLDHVRENRRGLSODLLNMCQIAGKMSYLEVR	840
Qy	841	LVHRDLARNVLVYKSPNNHYKITDFGLARLLIDETEHADGKVPIMKMALESILRRFT	900
Db	841	LVHRDLARNVLVYKSPNNHYKITDFGLARLLIDETEHADGKVPIMKMALESILRRFT	900
Qy	901	HOSDWSVGVTVWELMTFGAKPYDGI PAEIPDLLEKGRLPQPICTIDVYMIWYKCM	960
Db	901	HOSDWSVGVTVWELMTFGAKPYDGI PAEIPDLLEKGRLPQPICTIDVYMIWYKCM	960
Qy	961	IDSECRPFRRELVSFESFMRARDPORFVYIQLNEDLPASPLDSTFYSRLLEDMDGLVDA	1020
Db	961	IDSECRPFRRELVSFESFMRARDPORFVYIQLNEDLPASPLDSTFYSRLLEDMDGLVDA	1020
Qy	1021	EEYLPQOGFCPPDAPAGAGVWHRHRSSSTRSGGDLTLGLAFSEEBAPRSLAPSEG	1080
Db	1021	EEYLPQOGFCPPDAPAGAGVWHRHRSSSTRSGGDLTLGLAFSEEBAPRSLAPSEG	1080
Qy	1081	AGSDVFDGDLGMAKGLQSLPTHDPSPLOQRSEDPVPLPSETTGVAAPLTCSPQPEVY	1140
Db	1081	AGSDVFDGDLGMAKGLQSLPTHDPSPLOQRSEDPVPLPSETTGVAAPLTCSPQPEVY	1140
Qy	1141	NOPDVRPOPSPSREBPLPAARPAAGATLESPKTLSPGKGVVDVAFGGAVENTPEYLPQ	1200
Db	1141	NOPDVRPOPSPSREBPLPAARPAAGATLESPKTLSPGKGVVDVAFGGAVENTPEYLPQ	1200
Qy	1201	GGAAPQPHPPAPFSAFDNLVYWDODPPERGAPSTFKGTPTAENPEYLGIDVPY	1255
Db	1201	GGAAPQPHPPAPFSAFDNLVYWDODPPERGAPSTFKGTPTAENPEYLGIDVPY	1255
RESULT 7			
US-10-418-027-3			
; Sequence 3, Application US/10418027			
; Publication No. US20030224467A1			
GENERAL INFORMATION:			
; APPLICANT: Osborne, C. Kent			
; APPLICANT: Schieff, Rachel			
; APPLICANT: Bardou, Valerie			
; APPLICANT: Hilsenbeck, Susan			
; APPLICANT: Clark, Gary			
; APPLICANT: Wong, Jiemln			
; APPLICANT: Chamness, Gary			
; APPLICANT: Hopp, Torsten			
; TITLE OF INVENTION: AIB 1 as a prognostic marker and predictor of endocrine therapy			
; TITLE OF INVENTION: resistance			
; FILE REFERENCE: HO-P02396US1			
; CURRENT APPLICATION NUMBER: US/10/418,027			
; PRIOR FILING DATE: 2003-04-17			
; PRIOR APPLICATION NUMBER: US 60/373,237			
; PRIOR FILING DATE: 2002-04-17			
; NUMBER OF SEQ ID NOS: 3			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 3			
; LENGTH: 1255			
; TYPE: PRT			
; ORGANISM: Human			
US-10-418-027-3			

Query Match	100.0%;	Score 6815;	DB 4;	Length 1255;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1255; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	I	MELALACRWGLLALLP	PGASTOVCTGTDKRLPASPEHMLRLHXGCGVOVQNL	60
Db	1	MELALACRWGLLALLP	PGASTOVCTGTDKRLPASPEHMLRLHXGCGVOVQNL	60
QY	61	ELTYLPTNASLSP	ODIOEVGVYLIHANOVRVPLQRLIRVGTOLFEDNYALAVLDNG	120
Db	61	ELTYLPTNASLSP	ODIOEVGVYLIHANOVRVPLQRLIRVGTOLFEDNYALAVLDNG	120
QY	121	DPLNNTTPVTGAS	PGGLRELOLRSLTEILKGGVLIORNPOLCYODTILMKOIFPKHNOILA	180
Db	121	DPLNNTTPVTGAS	PGGLRELOLRSLTEILKGGVLIORNPOLCYODTILMKOIFPKHNOILA	180
QY	181	LTLIDTNRSRACHPC	SPCKSGRCKMGSSEBDCOSLTRVCAAGGRCCKGRLPTDCCHBQC	240
Db	181	LTLIDTNRSRACHPC	SPCKSGRCKMGSSEBDCOSLTRVCAAGGRCCKGRLPTDCCHBQC	240
QY	241	AAGCTGPHGSDCLAC	LHNHSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVAACP	300
Db	241	AAGCTGPHGSDCLAC	LHNHSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVAACP	300
QY	301	YNYISTDVGSCTLYC	PLHNOEVTADGTOREKCKSPCARCYGIGMHLEAVAYLSAN	360
Db	301	YNYISTDVGSCTLYC	PLHNOEVTADGTOREKCKSPCARCYGIGMHLEAVAYLSAN	360
QY	361	IOEPAGCKKIFGSLA	FPESFDGPASNTALPQEPOLVFETLEBITGYLISAMPDILP	420
Db	361	IOEPAGCKKIFGSLA	FPESFDGPASNTALPQEPOLVFETLEBITGYLISAMPDILP	420
QY	421	DLASFQNLQVIRGR	ILHNGAVSLTLQGLISBWLGRSLRELGSGIALIHNHTHLCFVHTV	480
Db	421	DLASFQNLQVIRGR	ILHNGAVSLTLQGLISBWLGRSLRELGSGIALIHNHTHLCFVHTV	480
QY	481	PWDOLFRRPHOALL	HTANRPREDECYGBELAHQICARHCKMGPRTQCVNCSQPLRGQEC	540
Db	481	PWDOLFRRPHOALL	HTANRPREDECYGBELAHQICARHCKMGPRTQCVNCSQPLRGQEC	540
QY	541	VEECRVLOGLRE	YVYNAHCLPCHRECPONGSVTCFEPBMDQCACAHYXDPFCVVARC	600
Db	541	VEECRVLOGLRE	YVYNAHCLPCHRECPONGSVTCFEPBMDQCACAHYXDPFCVVARC	600
QY	601	PSGVKPDLSYMP	IKWFPDEBEGACOPCPINCHSCYDLDKCGPAEORASPLTISIAYVG	660
Db	601	PSGVKPDLSYMP	IKWFPDEBEGACOPCPINCHSCYDLDKCGPAEORASPLTISIAYVG	660
QY	661	ILLVYVGVN	GGIILKRRQOKIRKTYMRLLQETELVEPLTPSGAMPQAOBRILKETEL	720
Db	661	ILLVYVGVN	GGIILKRRQOKIRKTYMRLLQETELVEPLTPSGAMPQAOBRILKETEL	720
QY	721	RKVVLGSGAG	FGTYVKGIMIPDGENVKIPVAIKVLRNTSPRANKBIIDEAYVMAGVSP	780
Db	721	RKVVLGSGAG	FGTYVKGIMIPDGENVKIPVAIKVLRNTSPRANKBIIDEAYVMAGVSP	780
QY	781	YVSRLLGICL	STYQVLQTLMPYGCILDHVRBNRGRLSODLLWCMQIAGKMSYLEBVR	840
Db	781	YVSRLLGICL	STYQVLQTLMPYGCILDHVRBNRGRLSODLLWCMQIAGKMSYLEBVR	840
QY	841	LVRHDLARBNLY	SPNHVKITTDGLARLLIDETRYADGKAYIKKMALESILRRFT	900
Db	841	LVRHDLARBNLY	SPNHVKITTDGLARLLIDETRYADGKAYIKKMALESILRRFT	900
QY	901	HOSDVM	SYGTVLWELMTFGAKPYDGI PARBI PDLEKEGRLPQPEICTIDVYMIWKCWM	960
Db	901	HOSDVM	SYGTVLWELMTFGAKPYDGI PARBI PDLEKEGRLPQPEICTIDVYMIWKCWM	960
QY	961	IDSECRPPFR	ELVSEFSRMAADPQAFVYIQNEDLGPAPPLDSTFRSLLEDDBDMDGLVDA	1020
Db	961	IDSECRPPFR	ELVSEFSRMAADPQAFVYIQNEDLGPAPPLDSTFRSLLEDDBDMDGLVDA	1020

Qy 1021 BEYLVPOQGFPCPDPAAGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPSEG 1080
Db 1021 BEYLVPOQGFPCPDPAAGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPSEG 1080
Qy 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYAPLTCSPOPEVY 1140
Db 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYAPLTCSPOPEVY 1140
Qy 1141 NOPDVRPOPSPREBGLPAARPAAGATLERPKTLSPGKGVVVDVAFGAVENPEYLTPQ 1200
Db 1141 NOPDVRPOPSPREBGLPAARPAAGATLERPKTLSPGKGVVVDVAFGAVENPEYLTPQ 1200
Qy 1201 GGAAPQHPHPPAFSPAFDNLYYWDDPPERGAPESTFKGTPTAENPEYLGLDVPV 1255
Db 1201 GGAAPQHPHPPAFSPAFDNLYYWDDPPERGAPESTFKGTPTAENPEYLGLDVPV 1255

RESULT 8
US-10-394-322A-17
; Sequence 17, Application US/10394322A
; Publication No. US20030232391A1
; GENERAL INFORMATION:
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.
; APPLICANT: PRESCOTT, JOHN C.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: 39750-0006 US
; CURRENT APPLICATION NUMBER: US/10/394,322A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/366,892
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FaetsSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-394-322A-17

Query Match 100.0%; Score 6815; DB 4; Length 1255;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MELALACRWGLLLALPFGAASGVCTGDMKRLPASPETHLDMRLHYOCQVVOGNTL 60
Db 1 MELALACRWGLLLALPFGAASGVCTGDMKRLPASPETHLDMRLHYOCQVVOGNTL 60
Qy 61 ELTYLPTNASLSFLQDIQEVGVYLIANQVRQVPLQRLIVRGTVLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIQEVGVYLIANQVRQVPLQRLIVRGTVLFEDNYALAVLDNG 120
Qy 121 DPLNNTPTPTGASPGGLRELQRLSLTELKGGVLIQRPOLCYOTIIMKDI FHKNNOLA 180
Db 121 DPLNNTPTPTGASPGGLRELQRLSLTELKGGVLIQRPOLCYOTIIMKDI FHKNNOLA 180
Qy 181 LTLIDTNSRACHPCSPMKGSRGWESSEBDCOSLTRTVACAGCARCKPLPTDCCHBQC 240
Db 181 LTLIDTNSRACHPCSPMKGSRGWESSEBDCOSLTRTVACAGCARCKPLPTDCCHBQC 240
Qy 241 AAGCTGPKJSDCLACLFHNSGICELHCPALVTYNTDTESMPNPEGRTYFGASCVTACP 300
Db 241 AAGCTGPKJSDCLACLFHNSGICELHCPALVTYNTDTESMPNPEGRTYFGASCVTACP 300
Qy 301 XNYLSTDVGSCTLVCPRLHNOVTAEDGTORCKRSPCARVCYGI GMEHLREVRVATGAN 360
Db 301 XNYLSTDVGSCTLVCPRLHNOVTAEDGTORCKRSPCARVCYGI GMEHLREVRVATGAN 360
Qy 361 IOEPAGCKKIFGSLAFLPESFGDDPASNTAPLQPEOLQVFETLBEITGTYLSAPDSLP 420
Db 361 IOEPAGCKKIFGSLAFLPESFGDDPASNTAPLQPEOLQVFETLBEITGTYLSAPDSLP 420
Qy 421 DLSVFQNLQVIRGRILLHNGAYSLTQGLGISWLGRLSRLSELGSGALIHNTHLCFVHTV 480
Db 421 DLSVFQNLQVIRGRILLHNGAYSLTQGLGISWLGRLSRLSELGSGALIHNTHLCFVHTV 480

Qy 481 PMDQLFRNPHQALLTANRPEDECEVGEGLACHQLCARGHCWGPBTQCVCNCSQFLRGEC 540
Db 481 PMDQLFRNPHQALLTANRPEDECEVGEGLACHQLCARGHCWGPBTQCVCNCSQFLRGEC 540
Qy 541 VEECEVLOGLPREYVNAHCLPCHEPCOPONGSVTCFGEADQCYACHYKDPPECVAC 600
Db 541 VEECEVLOGLPREYVNAHCLPCHEPCOPONGSVTCFGEADQCYACHYKDPPECVAC 600
Qy 601 PSGVCPDLSYMPIMKPEPBEBCACOPCPINCHSCVDLDDKGPBARORASPLTSSISAVYG 660
Db 601 PSGVCPDLSYMPIMKPEPBEBCACOPCPINCHSCVDLDDKGPBARORASPLTSSISAVYG 660
Qy 661 ILVVVLGVVFGLIKRRQOKIRKTYMRLLOETELVLEPLTPSGAMPNOQMRILKETEL 720
Db 661 ILVVVLGVVFGLIKRRQOKIRKTYMRLLOETELVLEPLTPSGAMPNOQMRILKETEL 720
Qy 721 RKVYLSGAGFTYKGIWIPDGENVKI PVAIKVLENTSPKANKELIDEAYVMAGVSP 780
Db 721 RKVYLSGAGFTYKGIWIPDGENVKI PVAIKVLENTSPKANKELIDEAYVMAGVSP 780
Qy 781 YVSRLLGICLSTVQVLVQMLPVGCLDHYRENRGLQSODLNMCMQIAKMSYLBDR 840
Db 781 YVSRLLGICLSTVQVLVQMLPVGCLDHYRENRGLQSODLNMCMQIAKMSYLBDR 840
Qy 841 LVHRDLAARNVLYKSPNHVKTDFGLARLLODETEYADGKVPKIMMALESILRRFT 900
Db 841 LVHRDLAARNVLYKSPNHVKTDFGLARLLODETEYADGKVPKIMMALESILRRFT 900
Qy 901 HQSDVSYGVTVWELMTFGAKPYDGI PAREIPDLLEKBERLPDPICTIDVYMIWKCWM 960
Db 901 HQSDVSYGVTVWELMTFGAKPYDGI PAREIPDLLEKBERLPDPICTIDVYMIWKCWM 960
Qy 961 IDSECRPRFRELVSFSSMAADPORFVYIQMEDLAPASPLDSTFRSLBEDMDGLVDA 1020
Db 961 IDSECRPRFRELVSFSSMAADPORFVYIQMEDLAPASPLDSTFRSLBEDMDGLVDA 1020
Qy 1021 BEYLVPOQGFPCPDPAAGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPSEG 1080
Db 1021 BEYLVPOQGFPCPDPAAGAGVHHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPSEG 1080
Qy 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYAPLTCSPOPEVY 1140
Db 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYAPLTCSPOPEVY 1140
Qy 1141 NOPDVRPOPSPREBGLPAARPAAGATLERPKTLSPGKGVVVDVAFGAVENPEYLTPQ 1200
Db 1141 NOPDVRPOPSPREBGLPAARPAAGATLERPKTLSPGKGVVVDVAFGAVENPEYLTPQ 1200
Qy 1201 GGAAPQHPHPPAFSPAFDNLYYWDDPPERGAPESTFKGTPTAENPEYLGLDVPV 1255
Db 1201 GGAAPQHPHPPAFSPAFDNLYYWDDPPERGAPESTFKGTPTAENPEYLGLDVPV 1255

RESULT 9
US-10-245-871-553
; Sequence 553, Application US/10245871
; Publication No. US20030235594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: 11-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 553
; LENGTH: 1255

TYPE: PRT
ORGANISM: Homo sapiens
US-10-245-871-553

Query Match 100.0%; Score 6815; DB 4; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MELALACRMGLLALLPFGAASVCTGTDMLRLPASETHLDMRLHYGCGVQGNL 60
DB 1 MELALACRMGLLALLPFGAASVCTGTDMLRLPASETHLDMRLHYGCGVQGNL 60
QY 61 ELTYLPTNASLSFLDIOEVQGYVLIANOVQVPLQRLRYRGTQLFEDNYALAVDNG 120
DB 61 ELTYLPTNASLSFLDIOEVQGYVLIANOVQVPLQRLRYRGTQLFEDNYALAVDNG 120
QY 121 DPLNNTTPTVGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNNOLA 180
DB 121 DPLNNTTPTVGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNNOLA 180
QY 181 LTLIDTNRBRACHPCSPCKSGRCWGSSEDCQSLTRTVCAAGGACRCKGPLPTDCHEQC 240
DB 181 LTLIDTNRBRACHPCSPCKSGRCWGSSEDCQSLTRTVCAAGGACRCKGPLPTDCHEQC 240
QY 241 AAGCTGPRKSDCLACLFHNSGICELHCPALVTYNTDTPESMPNEGRTTGASCVTACP 300
DB 241 AAGCTGPRKSDCLACLFHNSGICELHCPALVTYNTDTPESMPNEGRTTGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPRLHNOEYTAEDGTORCEKSKPCARVYGI GMEHLRVRVTSAN 360
DB 301 YNYLSTDVGSCTLVCPRLHNOEYTAEDGTORCEKSKPCARVYGI GMEHLRVRVTSAN 360
QY 361 IOEPAGCKKIFGSLAFLESPFDGASNTAPLQPPOLQFETLKEITGYLVI SAMPDILP 420
DB 361 IOEPAGCKKIFGSLAFLESPFDGASNTAPLQPPOLQFETLKEITGYLVI SAMPDILP 420
QY 421 DLVSFQNLQVIRGRILHNGAVSLTQGLGISWLGRLSRLSGLALIHNNTHLCFVATV 480
DB 421 DLVSFQNLQVIRGRILHNGAVSLTQGLGISWLGRLSRLSGLALIHNNTHLCFVATV 480
QY 481 PWDQLFRNPQALHTTANRPEDECVBEGLAGHQLCARHGCWGPRTQCYNCSQFLRGEC 540
DB 481 PWDQLFRNPQALHTTANRPEDECVBEGLAGHQLCARHGCWGPRTQCYNCSQFLRGEC 540
QY 541 VEECGVLQGLPREYVNAHCLPCHPECOPONGSVTCFGBADQVCAAHYKDPFCVARC 600
DB 541 VEECGVLQGLPREYVNAHCLPCHPECOPONGSVTCFGBADQVCAAHYKDPFCVARC 600
QY 601 PSGVPRDLSSYMPIMKFPDEBGAQCPINCTHSCVDLDDKGCAPAEQASPLTSII SAAVVG 660
DB 601 PSGVPRDLSSYMPIMKFPDEBGAQCPINCTHSCVDLDDKGCAPAEQASPLTSII SAAVVG 660
QY 661 ILLVVLGVVFGIILIKRROQIKRTYMRRLLOETELVEPLTSGAMPNOAQMRILKETEL 720
DB 661 ILLVVLGVVFGIILIKRROQIKRTYMRRLLOETELVEPLTSGAMPNOAQMRILKETEL 720
QY 721 RVKVLGSGAGFTYVKGIMIPGENYKIPVAIKVLRENTSPKANEIIDEAVMAGVSP 780
DB 721 RVKVLGSGAGFTYVKGIMIPGENYKIPVAIKVLRENTSPKANEIIDEAVMAGVSP 780
QY 781 YVSRLLGICLTSTVQLVQLMPEYGLLDHVRNENRGLSGODLLANCMQIAKMSYLEDVR 840
DB 781 YVSRLLGICLTSTVQLVQLMPEYGLLDHVRNENRGLSGODLLANCMQIAKMSYLEDVR 840
QY 841 LVHRDLAARNVLKSPNHYKITDFGLARLLIDETBYHADGSKVPIKMMALESILRRFT 900
DB 841 LVHRDLAARNVLKSPNHYKITDFGLARLLIDETBYHADGSKVPIKMMALESILRRFT 900
QY 901 HOSDWASVGVTVMEIETFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVMIIVKCM 960
DB 901 HOSDWASVGVTVMEIETFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVMIIVKCM 960
QY 961 IDSECRPRFRELVSERFARMADPQRFVLIQNEIDLGPASPLDSTFYRSLLEDDMDGLVDA 1020
```

```
DB 961 IDSECRPRFRELVSERFARMADPQRFVLIQNEIDLGPASPLDSTFYRSLLEDDMDGLVDA 1020
QY 1021 BEYLVPQGFPCPPAPAGAGMWHRRSSSTRSGGDLTLGLEPSEBAPRSLAPSE 1080
DB 1021 BEYLVPQGFPCPPAPAGAGMWHRRSSSTRSGGDLTLGLEPSEBAPRSLAPSE 1080
QY 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLOKRYSEDPTVPLPSEITDGYVAPLTCSPOPEYV 1140
DB 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLOKRYSEDPTVPLPSEITDGYVAPLTCSPOPEYV 1140
QY 1141 NOPDVRQPPSPRRGEPPLAARPAATLERPKTSLPGKNGVYKVYVARGAVNEBEYLTPQ 1200
DB 1141 NOPDVRQPPSPRRGEPPLAARPAATLERPKTSLPGKNGVYKVYVARGAVNEBEYLTPQ 1200
QY 1201 GGAAPQHPHPPAPAPADNLVYWDQDPERGAPSTKGTAEANPEYLGIDVPE 1255
DB 1201 GGAAPQHPHPPAPAPADNLVYWDQDPERGAPSTKGTAEANPEYLGIDVPE 1255
```

RESULT 10

```
US-10-469-162-3
; Sequence 3, Application US/10469162
; Publication No. US20040052811A1
; GENERAL INFORMATION:
; APPLICANT: Ziefelski, Christoph
; APPLICANT: Behamberger, Hubert
; APPLICANT: Breiteneder, Helmo
; APPLICANT: Jensen-Jarolim, Erika
; APPLICANT: Scheiner, Otto
; TITLE OF INVENTION: Vaccines Against Cancerous Diseases Associated With the HER-2/neu
; FILE REFERENCE: K 38 132/3yv
; CURRENT APPLICATION NUMBER: US/10/469,162
; PRIOR FILING DATE: 2003-08-27
; PRIOR APPLICATION NUMBER: PCT/EP02/02111
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: EP 01104943.4
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(675)
; OTHER INFORMATION: Extracellular Domain
US-10-469-162-3
```

Query Match 100.0%; Score 6815; DB 4; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MELALACRMGLLALLPFGAASVCTGTDMLRLPASETHLDMRLHYGCGVQGNL 60
DB 1 MELALACRMGLLALLPFGAASVCTGTDMLRLPASETHLDMRLHYGCGVQGNL 60
QY 61 ELTYLPTNASLSFLDIOEVQGYVLIANOVQVPLQRLRYRGTQLFEDNYALAVDNG 120
DB 61 ELTYLPTNASLSFLDIOEVQGYVLIANOVQVPLQRLRYRGTQLFEDNYALAVDNG 120
QY 121 DPLNNTTPTVGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNNOLA 180
DB 121 DPLNNTTPTVGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNNOLA 180
QY 181 LTLIDTNRBRACHPCSPCKSGRCWGSSEDCQSLTRTVCAAGGACRCKGPLPTDCHEQC 240
DB 181 LTLIDTNRBRACHPCSPCKSGRCWGSSEDCQSLTRTVCAAGGACRCKGPLPTDCHEQC 240
QY 241 AAGCTGPRKSDCLACLFHNSGICELHCPALVTYNTDTPESMPNEGRTTGASCVTACP 300
```

```

Db 241 AAGCTGPKHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNEGRTTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRATSAN 360
Db 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRATSAN 360
Qy 361 IOEPAGCKKIFGSLAFLPESPDGDPASNTAPLQPEQLQVFETLBEITGLYISAMPDSL 420
Db 361 IOEPAGCKKIFGSLAFLPESPDGDPASNTAPLQPEQLQVFETLBEITGLYISAMPDSL 420
Qy 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISWLGRLSRLSGLALIHNTHLCFVHTV 480
Db 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISWLGRLSRLSGLALIHNTHLCFVHTV 480
Qy 481 PWDOLFRRNHQALLHTANRPEDECVBEGLAGHQLCARHGWCPGPQCNCGQPLRGQEC 540
Db 481 PWDOLFRRNHQALLHTANRPEDECVBEGLAGHQLCARHGWCPGPQCNCGQPLRGQEC 540
Qy 541 VBECHVLOGLPREYVNAHCLPCHPECOPONGSVTCFGBEADQCVACHYKDPPECVARN 600
Db 541 VBECHVLOGLPREYVNAHCLPCHPECOPONGSVTCFGBEADQCVACHYKDPPECVARN 600
Qy 601 PSGVPRDLSYMPIMKFPDEBGAQCPCPINCTHSCVDLDDKGCPRABORASPLTSIIISAV 660
Db 601 PSGVPRDLSYMPIMKFPDEBGAQCPCPINCTHSCVDLDDKGCPRABORASPLTSIIISAV 660
Qy 661 ILVVVVGAVFGIILKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
Db 661 ILVVVVGAVFGIILKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
Qy 721 RKVKVLSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPRANKIILDEAYVMAGVGP 780
Db 721 RKVKVLSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPRANKIILDEAYVMAGVGP 780
Qy 781 YVSRLLGICLSTSTVOLVTQLMFYGCLDHYRENRGLSQDILLNMQIJAQMSYLEDR 840
Db 781 YVSRLLGICLSTSTVOLVTQLMFYGCLDHYRENRGLSQDILLNMQIJAQMSYLEDR 840
Qy 841 LVHRLLAARNVLKSPNHVKITDFGLARLLDIDETRYHADGKVPKXMALESILRRRT 900
Db 841 LVHRLLAARNVLKSPNHVKITDFGLARLLDIDETRYHADGKVPKXMALESILRRRT 900
Qy 901 HOSDWVSVGVTWELMTFGAKPYDGI PARBIPLLEKGERLPORPICTIDVTMIMVKCM 960
Db 901 HOSDWVSVGVTWELMTFGAKPYDGI PARBIPLLEKGERLPORPICTIDVTMIMVKCM 960
Qy 961 IDSECRPRELVSERSMARDPQRFVVIQNEBLAGPASPDLSTFYRSLLEDDMGDLVDA 1020
Db 961 IDSECRPRELVSERSMARDPQRFVVIQNEBLAGPASPDLSTFYRSLLEDDMGDLVDA 1020
Qy 1021 BEYLVPOQGFPCPDPAFGAGVHHRHSSSTRSGGDLTTLGBSEBEAPRSPPLAPSG 1080
Db 1021 BEYLVPOQGFPCPDPAFGAGVHHRHSSSTRSGGDLTTLGBSEBEAPRSPPLAPSG 1080
Qy 1081 AGSDVFDGDLGGAAGLQSLPTHPSPLOKISEPTVPLPSETDGYAVPLTCSQPEV 1140
Db 1081 AGSDVFDGDLGGAAGLQSLPTHPSPLOKISEPTVPLPSETDGYAVPLTCSQPEV 1140
Qy 1141 NOPDVRPOPSPREBGLPARAPAGATLERPKTILSPKINGVADKAPFGAVENPEYLTPQ 1200
Db 1141 NOPDVRPOPSPREBGLPARAPAGATLERPKTILSPKINGVADKAPFGAVENPEYLTPQ 1200
Qy 1201 GGAAQPHPPAFSPAFDNLTYWDDPPRRGAPSPSTFKGTPTAENPEYLGDLVP 1255
Db 1201 GGAAQPHPPAFSPAFDNLTYWDDPPRRGAPSPSTFKGTPTAENPEYLGDLVP 1255

```

```

; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: I1-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 553
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-253-286-553

Query Match 100.0%; Score 6815; DB 4; Length 1255;
Best Local Similarity 100.0%; Pred. No; 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 METALCRWGILLALLPFGAASSTOVCSTDMKLRPASPTLMDLRHLXGCGVVGVL 60
Db 1 METALCRWGILLALLPFGAASSTOVCSTDMKLRPASPTLMDLRHLXGCGVVGVL 60
Qy 61 ELTYLPTNASLFLQDIOEVQGYVLIAHNVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASLFLQDIOEVQGYVLIAHNVQVPLQRLIRVGTQLFEDNYALAVLDNG 120
Qy 121 DPLNNTPTVAGSPGLRELOLRSLTEILKGGVLIQRNPOLCYQDTILMKDIFHKNOLA 180
Db 121 DPLNNTPTVAGSPGLRELOLRSLTEILKGGVLIQRNPOLCYQDTILMKDIFHKNOLA 180
Qy 181 LTLIDTPNSRACHPSSPMCKSGRCWGESSBDQSLTRTVACGGCARCGPLPTDCHEQC 240
Db 181 LTLIDTPNSRACHPSSPMCKSGRCWGESSBDQSLTRTVACGGCARCGPLPTDCHEQC 240
Qy 241 AAGCTGPKHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNEGRTTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACIHFHNSGICELHCPALVTYNTDFESMPNEGRTTFGASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRATSAN 360
Db 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVRATSAN 360
Qy 361 IOEPAGCKKIFGSLAFLPESPDGDPASNTAPLQPEQLQVFETLBEITGLYISAMPDSL 420
Db 361 IOEPAGCKKIFGSLAFLPESPDGDPASNTAPLQPEQLQVFETLBEITGLYISAMPDSL 420
Qy 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISWLGRLSRLSGLALIHNTHLCFVHTV 480
Db 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISWLGRLSRLSGLALIHNTHLCFVHTV 480
Qy 481 PWDOLFRRNHQALLHTANRPEDECVBEGLAGHQLCARHGWCPGPQCNCGQPLRGQEC 540
Db 481 PWDOLFRRNHQALLHTANRPEDECVBEGLAGHQLCARHGWCPGPQCNCGQPLRGQEC 540
Qy 541 VBECHVLOGLPREYVNAHCLPCHPECOPONGSVTCFGBEADQCVACHYKDPPECVARN 600
Db 541 VBECHVLOGLPREYVNAHCLPCHPECOPONGSVTCFGBEADQCVACHYKDPPECVARN 600
Qy 601 PSGVPRDLSYMPIMKFPDEBGAQCPCPINCTHSCVDLDDKGCPRABORASPLTSIIISAV 660
Db 601 PSGVPRDLSYMPIMKFPDEBGAQCPCPINCTHSCVDLDDKGCPRABORASPLTSIIISAV 660
Qy 661 ILVVVVGAVFGIILKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
Db 661 ILVVVVGAVFGIILKRRQOKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
Qy 721 RKVKVLSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPRANKIILDEAYVMAGVGP 780
Db 721 RKVKVLSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPRANKIILDEAYVMAGVGP 780

```

```

RESULT 11
US-10-253-286-553
; Sequence 553, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT

```

Qy 781 YVSRLLGICLSTVQLVTLQMPYGCGLDHHVRENGRLGSODLLNMCQIAGKMSYLEBVR 840
Db 781 YVSRLLGICLSTVQLVTLQMPYGCGLDHHVRENGRLGSODLLNMCQIAGKMSYLEBVR 840
Qy 841 LVHRDLAARNLVKSPNHNKTTDFGLARLLDIDETEVHADGKVPKIMMALESILRRFT 900
Db 841 LVHRDLAARNLVKSPNHNKTTDFGLARLLDIDETEVHADGKVPKIMMALESILRRFT 900
Qy 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITIDVYMIWVKCM 960
Db 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITIDVYMIWVKCM 960
Qy 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGPAAPLDSTFYRSLLEDMDGLVDA 1020
Db 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGPAAPLDSTFYRSLLEDMDGLVDA 1020
Qy 1021 EBYLVPQGGFPCPDPAFGAGVHHRHSSSTRSGGDLTLGLPSEEBAPRSLAPSEG 1080
Db 1021 EBYLVPQGGFPCPDPAFGAGVHHRHSSSTRSGGDLTLGLPSEEBAPRSLAPSEG 1080
Qy 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQRYSDPTVPLPSETDGYVAPLTCSFPQRYV 1140
Db 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQRYSDPTVPLPSETDGYVAPLTCSFPQRYV 1140
Qy 1141 NOPDVRPOPSPRSGPLPAARPAAGATLBRPKTLSPGKGVVDVAFGAVENPEYLTPO 1200
Db 1141 NOPDVRPOPSPRSGPLPAARPAAGATLBRPKTLSPGKGVVDVAFGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPFAPFADNLVYWDQDPPBRGAPSTFTGTPAENPEYLGIDVPV 1255
Db 1201 GGAAPQHPHPAPFAPFADNLVYWDQDPPBRGAPSTFTGTPAENPEYLGIDVPV 1255

RESULT 12

US-10-762-128-6
; Sequence 6, Application US/10762128
; Publication No. US20040219161A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Diets, Mary L.
; APPLICANT: Hellstrom, Inggerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409C1
; CURRENT APPLICATION NUMBER: US/10/762.128
; PRIOR FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US 09/441.411
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-762-128-6

Query Match 100.0%; Score 6815; DB 5; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MELALACWGLLALLPFGAASVCTGTDMKLRIPASPEHTLMDLRHLYGSCQVYQCNL 60
Db 1 MELALACWGLLALLPFGAASVCTGTDMKLRIPASPEHTLMDLRHLYGSCQVYQCNL 60
Qy 61 ELTYLPTNASLSFLDIDDEVQGYVLIANQVQVPLQRLRIVRGTLFEDNVALAVLDNG 120
Db 61 ELTYLPTNASLSFLDIDDEVQGYVLIANQVQVPLQRLRIVRGTLFEDNVALAVLDNG 120
Qy 121 DPLNNTTVYTGASPGGLRELQRLSTELIKGGVLIQRPOLCYQDTIIMKQIFHKNQOLA 180
Db 121 DPLNNTTVYTGASPGGLRELQRLSTELIKGGVLIQRPOLCYQDTIIMKQIFHKNQOLA 180
Qy 181 LTLIDTNRSRACHPCSPCKGSRCWGESSEDCQSLJRTVCAAGCARCKGPLPTDCHEQC 240

Db 181 LTLIDTNRSRACHPCSPCKGSRCWGESSEDCQSLJRTVCAAGCARCKGPLPTDCHEQC 240
Qy 241 AAGCTGPKHSDCIACLPNHSIGICELHCPALVTYNTPTFSMPNBEGRYTGASCVTACP 300
Db 241 AAGCTGPKHSDCIACLPNHSIGICELHCPALVTYNTPTFSMPNBEGRYTGASCVTACP 300
Qy 301 YNVLSTVGSCTLYVCPHNOEVTAEADGTORCEKSKCARVCYGLGMEHLREVAVTSAN 360
Db 301 YNVLSTVGSCTLYVCPHNOEVTAEADGTORCEKSKCARVCYGLGMEHLREVAVTSAN 360
Qy 361 IOEFAGCKKIFGSLAPLPESFDGPASNTAPLOEBOLOVEFTLEITGYLYISAMPDLP 420
Db 361 IOEFAGCKKIFGSLAPLPESFDGPASNTAPLOEBOLOVEFTLEITGYLYISAMPDLP 420
Qy 421 DLSVFONLQVIRGILHNGAYSLTLOGIGISMLGRSLRELSGLALIHNTHLCFVYTV 480
Db 421 DLSVFONLQVIRGILHNGAYSLTLOGIGISMLGRSLRELSGLALIHNTHLCFVYTV 480
Qy 481 PMDQLFRNPQALHTANRPEDECVGEGACLOLCARGHCWGPPTQCVCNSQFLRGQC 540
Db 481 PMDQLFRNPQALHTANRPEDECVGEGACLOLCARGHCWGPPTQCVCNSQFLRGQC 540
Qy 541 VEBECRYLQGLPREVYNAHCLPCHPECOQNGSYTCGFRADOCVACAHYDPPFCVARC 600
Db 541 VEBECRYLQGLPREVYNAHCLPCHPECOQNGSYTCGFRADOCVACAHYDPPFCVARC 600
Qy 601 PSGVKPDLSTYMPIMKFPDEBACOPCPINCTHSCVDDDKCPCAPORASPLTISIAYVG 660
Db 601 PSGVKPDLSTYMPIMKFPDEBACOPCPINCTHSCVDDDKCPCAPORASPLTISIAYVG 660
Qy 661 ILLVVLGVVFGILLIKRQOKIRKRYMRLLQETELVEPLTPSGAMPNQAQRIKTEL 720
Db 661 ILLVVLGVVFGILLIKRQOKIRKRYMRLLQETELVEPLTPSGAMPNQAQRIKTEL 720
Qy 721 RKXVYLSSGAGTYTKGIVIPDGENVKIPVAIKTLRENTSKANKETILDEAYVMAAGVSP 780
Db 721 RKXVYLSSGAGTYTKGIVIPDGENVKIPVAIKTLRENTSKANKETILDEAYVMAAGVSP 780
Qy 781 YVSRLLGICLSTVQLVTLQMPYGCGLDHHVRENGRLGSODLLNMCQIAGKMSYLEBVR 840
Db 781 YVSRLLGICLSTVQLVTLQMPYGCGLDHHVRENGRLGSODLLNMCQIAGKMSYLEBVR 840
Qy 841 LVHRDLAARNLVKSPNHNKTTDFGLARLLDIDETEVHADGKVPKIMMALESILRRFT 900
Db 841 LVHRDLAARNLVKSPNHNKTTDFGLARLLDIDETEVHADGKVPKIMMALESILRRFT 900
Qy 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITIDVYMIWVKCM 960
Db 901 HOSDWSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITIDVYMIWVKCM 960
Qy 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGPAAPLDSTFYRSLLEDMDGLVDA 1020
Db 961 IDSECRPRFRELVSFSSRMARDPQRFVYIQNEDLGPAAPLDSTFYRSLLEDMDGLVDA 1020
Qy 1021 EBYLVPQGGFPCPDPAFGAGVHHRHSSSTRSGGDLTLGLPSEEBAPRSLAPSEG 1080
Db 1021 EBYLVPQGGFPCPDPAFGAGVHHRHSSSTRSGGDLTLGLPSEEBAPRSLAPSEG 1080
Qy 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQRYSDPTVPLPSETDGYVAPLTCSFPQRYV 1140
Db 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQRYSDPTVPLPSETDGYVAPLTCSFPQRYV 1140
Qy 1141 NOPDVRPOPSPRSGPLPAARPAAGATLBRPKTLSPGKGVVDVAFGAVENPEYLTPO 1200
Db 1141 NOPDVRPOPSPRSGPLPAARPAAGATLBRPKTLSPGKGVVDVAFGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPFAPFADNLVYWDQDPPBRGAPSTFTGTPAENPEYLGIDVPV 1255
Db 1201 GGAAPQHPHPAPFAPFADNLVYWDQDPPBRGAPSTFTGTPAENPEYLGIDVPV 1255

RESULT 13


```

US-10-484-067-1
; Sequence 1, Application US/10484067
; Publication No. US2004024166A1
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF CALIFORNIA
; APPLICANT: NELSON, Edward L.
; TITLE OF INVENTION: HER2/NEU TARGET ANTIGEN AND USE OF SAME TO STIMULATE AN IMMUNE RE
; FILE REFERENCE: UCI1170-1
; CURRENT APPLICATION NUMBER: US/10/484,067
; PRIOR FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: PCT/US02/22975
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 60/306,250
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent version 3.1
; SEQ ID NO: 1
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-484-067-1

Query Match      100.0%; Score 6815; DB 5; Length 1255;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MELALACWGLLALLPPGAASVCTGTDMLRLPASPEHLDMRLHYOGCCVVGNTL 60
DB      1 MELALACWGLLALLPPGAASVCTGTDMLRLPASPEHLDMRLHYOGCCVVGNTL 60
QY      61 ELYLPTNASISFLDIOEVOGYVLIANOVQVPLQRLRIYRGTLFEDNYALAVLNG 120
DB      61 ELYLPTNASISFLDIOEVOGYVLIANOVQVPLQRLRIYRGTLFEDNYALAVLNG 120
QY      121 DPLNNTTPTVGTASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNOQLA 180
DB      121 DPLNNTTPTVGTASPGGLRELOLRSLTEILKGGVLIQRNPOLCYODTIIMKDI FHKNOQLA 180
QY      181 LTLIDTNSRACHPSSPMCKSGRCSESEDCOSLTRVYAGGACRCGRLPTDCCHQC 240
DB      181 LTLIDTNSRACHPSSPMCKSGRCSESEDCOSLTRVYAGGACRCGRLPTDCCHQC 240
QY      241 AAGCTGPKHSDDLACLFHNSGICELHCPALVTYNTDFEESPNPEGRTTFGASCVTACP 300
DB      241 AAGCTGPKHSDDLACLFHNSGICELHCPALVTYNTDFEESPNPEGRTTFGASCVTACP 300
QY      301 YNYLSTDVGSCTLVCPRLHNOEVTAEADGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
DB      301 YNYLSTDVGSCTLVCPRLHNOEVTAEADGTQRCCKSKPCARVCYGLGMEHLREVRVTSAN 360
QY      361 IOEPAGCKKIFGSLAFPLBESFGDPAASNTAPLOPEQLOVFETLEBTGLYLSAMPDSL 420
DB      361 IOEPAGCKKIFGSLAFPLBESFGDPAASNTAPLOPEQLOVFETLEBTGLYLSAMPDSL 420
QY      421 DLSVQNLQVTRGRILHNGAYSLTLOGLISWLGRLRELSGSLALHNNTHLCFVTV 480
DB      421 DLSVQNLQVTRGRILHNGAYSLTLOGLISWLGRLRELSGSLALHNNTHLCFVTV 480
QY      481 FMDQLFRNPHQALLHTANRPEDECVGEGLACHQLCARGHGWPGPTOCVNCSQLRGQEC 540
DB      481 FMDQLFRNPHQALLHTANRPEDECVGEGLACHQLCARGHGWPGPTOCVNCSQLRGQEC 540
QY      541 VEECVNLQGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPPECVARC 600
DB      541 VEECVNLQGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPPECVARC 600
QY      601 PSGVPRDLSYMPIMKFPBEGACOPCPINCTHSCVDLDDKGPABORASPLTSISAYVG 660
DB      601 PSGVPRDLSYMPIMKFPBEGACOPCPINCTHSCVDLDDKGPABORASPLTSISAYVG 660
QY      661 ILLVVLGVVFGILLIKRQOKIRKTYMRLLOETELVEPLTPSGAMPNOQMRILKETEL 720
DB      661 ILLVVLGVVFGILLIKRQOKIRKTYMRLLOETELVEPLTPSGAMPNOQMRILKETEL 720

```

```

QY      721 RKVKVLGSGAFGTVYKGIWIPDGENVKIPVALKVLRENTSPKANKIILDEAYVAGVGP 780
DB      721 RKVKVLGSGAFGTVYKGIWIPDGENVKIPVALKVLRENTSPKANKIILDEAYVAGVGP 780
QY      781 YVSRLLGICLTSTVOLVQLMPYGLLDHVRNRRGLSQDILLANCMQIAKMSYLEDVR 840
DB      781 YVSRLLGICLTSTVOLVQLMPYGLLDHVRNRRGLSQDILLANCMQIAKMSYLEDVR 840
QY      841 LVHRDLAARNVLKSPNNVKTIDREGLALLDIDETRYHADGKVPILKMALESILRRRT 900
DB      841 LVHRDLAARNVLKSPNNVKTIDREGLALLDIDETRYHADGKVPILKMALESILRRRT 900
QY      901 HOSDVSYGVTWELMTFGAKPYDGI PAREIPDLLEKGRPLPQPICTIDVYIMVCKM 960
DB      901 HOSDVSYGVTWELMTFGAKPYDGI PAREIPDLLEKGRPLPQPICTIDVYIMVCKM 960
QY      961 IDSECRPRFRELVSFSRWARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDMDGLVDA 1020
DB      961 IDSECRPRFRELVSFSRWARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDMDGLVDA 1020
QY      1021 ERYLVPQGFPCPDPAFGAGVHHRSSSTRSGGDLTGLBSEEBEAPSPILAPSRG 1080
DB      1021 ERYLVPQGFPCPDPAFGAGVHHRSSSTRSGGDLTGLBSEEBEAPSPILAPSRG 1080
QY      1081 AGSDVFDDGLGKAGKGLSLPTHDPSPLQRYSEDPVPLPSETGYVAPLTCSPQPEYV 1140
DB      1081 AGSDVFDDGLGKAGKGLSLPTHDPSPLQRYSEDPVPLPSETGYVAPLTCSPQPEYV 1140
QY      1141 NOPDVRPOPSPBREGPLPAARPAATLERPKTLSPGKGVVVDVAFGAVENPEYLTPO 1200
DB      1141 NOPDVRPOPSPBREGPLPAARPAATLERPKTLSPGKGVVVDVAFGAVENPEYLTPO 1200
QY      1201 GGAAPQHPPPAFSPFNLYWDDPPBKGAPSTFKGTFAENPEYLGDLV 1255
DB      1201 GGAAPQHPPPAFSPFNLYWDDPPBKGAPSTFKGTFAENPEYLGDLV 1255

RESULT 14
US-10-723-860-9
; Sequence 9, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natsaba
; APPLICANT: Ginsburg, Wendy M.
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882, 0193, NUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patent version 3.2
; SEQ ID NO: 9
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-9

Query Match      100.0%; Score 6815; DB 5; Length 1255;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MELALACWGLLALLPPGAASVCTGTDMLRLPASPEHLDMRLHYOGCCVVGNTL 60
DB      1 MELALACWGLLALLPPGAASVCTGTDMLRLPASPEHLDMRLHYOGCCVVGNTL 60
QY      61 ELYLPTNASISFLDIOEVOGYVLIANOVQVPLQRLRIYRGTLFEDNYALAVLNG 120
DB      61 ELYLPTNASISFLDIOEVOGYVLIANOVQVPLQRLRIYRGTLFEDNYALAVLNG 120

```


QY 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKKGVLIOBNPOLCYODTIIMKDI FHKNNOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKKGVLIOBNPOLCYODTIIMKDI FHKNNOLA 180
QY 181 LTLIDTNSRACHPCSPCKSGRCWGESSESDCOSLTRVCGAGGRCARCGPLPTDCCHQC 240
DB 181 LTLIDTNSRACHPCSPCKSGRCWGESSESDCOSLTRVCGAGGRCARCGPLPTDCCHQC 240
QY 241 AAGCTGPGHSDCLACLHFNHSGICELHCPALVTYNTDFFESMPNBEGRYTFGASCVTACP 300
DB 241 AAGCTGPGHSDCLACLHFNHSGICELHCPALVTYNTDFFESMPNBEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREBRAVTSAN 360
DB 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREBRAVTSAN 360
QY 361 IOEPFAGCKKIFGSLAFPLESPFDGDSANTAPLOPQLOVFETLEBITGYLYISAMPDSL 420
DB 361 IOEPFAGCKKIFGSLAFPLESPFDGDSANTAPLOPQLOVFETLEBITGYLYISAMPDSL 420
QY 421 DLSVFQNLQVIRGRILHNGAYSILTQGLIGISWLGRLSRLBSGLALJHNHTLCEFYATV 480
DB 421 DLSVFQNLQVIRGRILHNGAYSILTQGLIGISWLGRLSRLBSGLALJHNHTLCEFYATV 480
QY 481 PWDOLFRRNPQALHTTANRPEDECVBEGLAGHOLCARHCWGPPTQCVCNCSQFLRGQEC 540
DB 481 PWDOLFRRNPQALHTTANRPEDECVBEGLAGHOLCARHCWGPPTQCVCNCSQFLRGQEC 540
QY 541 VEECKVLOGLPREYVNAHCLPCHBECOPONGSVTCFGBEADQCAAHYDOPPCVARC 600
DB 541 VEECKVLOGLPREYVNAHCLPCHBECOPONGSVTCFGBEADQCAAHYDOPPCVARC 600
QY 601 PSGVPRDLSYMPIMKFPDEBAGCQPCPINCTHSCVDLDDKGPABQASPLTISIISAVVG 660
DB 601 PSGVPRDLSYMPIMKFPDEBAGCQPCPINCTHSCVDLDDKGPABQASPLTISIISAVVG 660
QY 661 ILVVVIGVFGILIKRRQOKIRKTYMRLLQETELVEPLTPSGAMPQOMRIKELTEL 720
DB 661 ILVVVIGVFGILIKRRQOKIRKTYMRLLQETELVEPLTPSGAMPQOMRIKELTEL 720
QY 721 RKVKVLGSGAGTGVKGIWIPDGENVKI PVALKVLRENTSPKANELLIDEAIVMAGVSP 780
DB 721 RKVKVLGSGAGTGVKGIWIPDGENVKI PVALKVLRENTSPKANELLIDEAIVMAGVSP 780
QY 781 YVSRLLGICLSTVOLVTQLMFYGCLLDHVENRGRLSQDLANCMOIAKMSLYLEDR 840
DB 781 YVSRLLGICLSTVOLVTQLMFYGCLLDHVENRGRLSQDLANCMOIAKMSLYLEDR 840
QY 841 LVHRDLAARNVLVKS PNHVKITDFGLARLLDIDETRYHADGGKVPDKMALESILRRFT 900
DB 841 LVHRDLAARNVLVKS PNHVKITDFGLARLLDIDETRYHADGGKVPDKMALESILRRFT 900
QY 901 HOSDVWSGYVYWEMLMTFGAKPYDGI PAEIPDLLEKBERLPQPICTIDVMIWVKCM 960
DB 901 HOSDVWSGYVYWEMLMTFGAKPYDGI PAEIPDLLEKBERLPQPICTIDVMIWVKCM 960
QY 961 IDSECRPFRELIVSEFSMARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDDMDLVDYA 1020
DB 961 IDSECRPFRELIVSEFSMARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDDMDLVDYA 1020
QY 1021 EBYLVPOQGFPCPDAPGAGWVHRHRSSTRSGGGDLTLGLEPSEEBAPRSLAPSBG 1080
DB 1021 EBYLVPOQGFPCPDAPGAGWVHRHRSSTRSGGGDLTLGLEPSEEBAPRSLAPSBG 1080
QY 1081 AGSDVPDDDLGMAKGLQSLPTHDPSLQRYSEBPTVPLPSETGYVAPLTCSQPEV 1140
DB 1081 AGSDVPDDDLGMAKGLQSLPTHDPSLQRYSEBPTVPLPSETGYVAPLTCSQPEV 1140
QY 1141 NOPDVRPOPSPBREGPLPAAPAGATLERPKTISPGKNGVADVAFAGGAVENPEYLTPO 1200
DB 1141 NOPDVRPOPSPBREGPLPAAPAGATLERPKTISPGKNGVADVAFAGGAVENPEYLTPO 1200
QY 1201 GGAAPQPHPPAPAFSPAFDNLVYWDODPPERGAPPESTFKGTPTAENPEYLGDLDPV 1255

DB 1201 GGAAPQPHPPAPAFSPAFDNLVYWDODPPERGAPPESTFKGTPTAENPEYLGDLDPV 1255
RESULT 15
US-10-871-708-9
; Sequence 9, Application US/10871708
; Publication No. US20050118186A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Chih-Sheng
; TITLE OF INVENTION: COMBINATIONS OF TUMOR-ASSOCIATED
; TITLE OF INVENTION: ANTIGENS IN COMPOSITIONS FOR VARIOUS TYPES OF CANCERS
; FILE REFERENCE: MANAK-035A
; CURRENT APPLICATION NUMBER: US/10/871,708
; PRIOR FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: 60/479,554
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Receptor protein-tyrosine kinase
US-10-871-708-9
Query Match 100.0%; Score 6815; DB 5; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELALCRWGLLALLPRGASSTQVCTGDMKRLPLPASPETHLDMLRHLVGGCOVQGNL 60
DB 1 MELALCRWGLLALLPRGASSTQVCTGDMKRLPLPASPETHLDMLRHLVGGCOVQGNL 60
QY 61 ELTYLPTNASLSFLQDIEVGGVYLIAHQVROYPLQRLIRVKTGFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSFLQDIEVGGVYLIAHQVROYPLQRLIRVKTGFEDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKKGVLIOBNPOLCYODTIIMKDI FHKNNOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELOLRSLTEILKKGVLIOBNPOLCYODTIIMKDI FHKNNOLA 180
QY 181 LTLIDTNSRACHPCSPCKSGRCWGESSESDCOSLTRVCGAGGRCARCGPLPTDCCHQC 240
DB 181 LTLIDTNSRACHPCSPCKSGRCWGESSESDCOSLTRVCGAGGRCARCGPLPTDCCHQC 240
QY 241 AAGCTGPGHSDCLACLHFNHSGICELHCPALVTYNTDFFESMPNBEGRYTFGASCVTACP 300
DB 241 AAGCTGPGHSDCLACLHFNHSGICELHCPALVTYNTDFFESMPNBEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREBRAVTSAN 360
DB 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREBRAVTSAN 360
QY 361 IOEPFAGCKKIFGSLAFPLESPFDGDSANTAPLOPQLOVFETLEBITGYLYISAMPDSL 420
DB 361 IOEPFAGCKKIFGSLAFPLESPFDGDSANTAPLOPQLOVFETLEBITGYLYISAMPDSL 420
QY 421 DLSVFQNLQVIRGRILHNGAYSILTQGLIGISWLGRLSRLBSGLALJHNHTLCEFYATV 480
DB 421 DLSVFQNLQVIRGRILHNGAYSILTQGLIGISWLGRLSRLBSGLALJHNHTLCEFYATV 480
QY 481 PWDOLFRRNPQALHTTANRPEDECVBEGLAGHOLCARHCWGPPTQCVCNCSQFLRGQEC 540
DB 481 PWDOLFRRNPQALHTTANRPEDECVBEGLAGHOLCARHCWGPPTQCVCNCSQFLRGQEC 540
QY 541 VEECKVLOGLPREYVNAHCLPCHBECOPONGSVTCFGBEADQCAAHYDOPPCVARC 600
DB 541 VEECKVLOGLPREYVNAHCLPCHBECOPONGSVTCFGBEADQCAAHYDOPPCVARC 600
QY 601 PSGVPRDLSYMPIMKFPDEBAGCQPCPINCTHSCVDLDDKGPABQASPLTISIISAVVG 660
DB 601 PSGVPRDLSYMPIMKFPDEBAGCQPCPINCTHSCVDLDDKGPABQASPLTISIISAVVG 660

Qy	661	ILLVVLGVFGLIKRROQKIRKTYMRRLLOETELVEPLTPSGAMPNOAQMRILKETEL	720
Db	661	ILVVVLGVFGLIKRROQKIRKTYMRRLLOETELVEPLTPSGAMPNOAQMRILKETEL	720
Qy	721	RKVVVLGSGARFTVYKGIWIPDGENVKI PVAIKVLRENTS PRANKELIDEAYVMAGVSP	780
Db	721	RKVVVLGSGARFTVYKGIWIPDGENVKI PVAIKVLRENTS PRANKELIDEAYVMAGVSP	780
Qy	781	YVSRLLGICLTSTVOLVQMPYGCLLDHVRENRGLSGODLLNWCQIAKMSYLEDVR	840
Db	781	YVSRLLGICLTSTVOLVQMPYGCLLDHVRENRGLSGODLLNWCQIAKMSYLEDVR	840
Qy	841	LVHRDLAARNVLVKS PNHVKITDPGLARLLDIDETEVHADGKVP IKWMALESILRRFT	900
Db	841	LVHRDLAARNVLVKS PNHVKITDPGLARLLDIDETEVHADGKVP IKWMALESILRRFT	900
Qy	901	HQSDVMSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVYMIWKCMM	960
Db	901	HQSDVMSYGVTVWELMTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVYMIWKCMM	960
Qy	961	IDSECRPRFRELVSFSPMARDPQRPVYIQNEDLGPA SPLDSTFYRSLLDDMDGLVDA	1020
Db	961	IDSECRPRFRELVSFSPMARDPQRPVYIQNEDLGPA SPLDSTFYRSLLDDMDGLVDA	1020
Qy	1021	EEYLVPOQGFPCPDPA PGAGVGHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPSEG	1080
Db	1021	EEYLVPOQGFPCPDPA PGAGVGHRRSSSTRSGGDLTLGLBPSBEEAPRSLAPSEG	1080
Qy	1081	AGSDVFDGDLGGAAGLQSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPQPEYV	1140
Db	1081	AGSDVFDGDLGGAAGLQSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPQPEYV	1140
Qy	1141	NQPDYRPOPPSPREGPLPAARPA GATLEBPXTLS PGKGVKDVAFAGAVENPEYLTPO	1200
Db	1141	NQPDYRPOPPSPREGPLPAARPA GATLEBPXTLS PGKGVKDVAFAGAVENPEYLTPO	1200
Qy	1201	GGAAPQPHPPPAFPAFNLVYWDODPPERGA PSTFKGTPTAENPEYLGLDVPV	1255
Db	1201	GGAAPQPHPPPAFPAFNLVYWDODPPERGA PSTFKGTPTAENPEYLGLDVPV	1255

Search completed: January 3, 2006, 11:31:45
Job time : 189 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2006, 11:08:37 ; Search time 50 Seconds
(without alignments)
2075.160 Million cell updates/sec

Title: US-09-930-125-2

Perfect score: 6815

Sequence: 1 METALCRWGLLALLPFGA.....TFKGTPTANPEYLGIDVIV 1255

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:

1: /cgn2_6/prodata/1/1aa/5 COMB.pep:*
2: /cgn2_6/prodata/1/1aa/6 COMB.pep:*
3: /cgn2_6/prodata/1/1aa/4 COMB.pep:*
4: /cgn2_6/prodata/1/1aa/ECTUS.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/RE.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6815	100.0	1255	US-08-625-101-2	Sequence 2, Appl1
2	6815	100.0	1255	US-08-356-786-2	Sequence 2, Appl1
3	6815	100.0	1255	US-09-441-411-6	Sequence 6, Appl1
4	6815	100.0	1255	US-09-167-516-2	Sequence 2, Appl1
5	6806	99.9	1255	US-08-467-083-68	Sequence 68, Appl1
6	6806	99.9	1255	US-08-414-417B-68	Sequence 68, Appl1
7	6806	99.9	1255	US-08-486-348A-68	Sequence 68, Appl1
8	6806	99.9	1255	US-08-468-545B-68	Sequence 68, Appl1
9	6806	99.9	1255	US-08-466-680B-68	Sequence 68, Appl1
10	6806	99.9	1255	US-09-527-487-2	Sequence 2, Appl1
11	6806	99.9	1255	US-09-811-115-3	Sequence 3, Appl1
12	6806	99.9	1255	US-09-354-533-68	Sequence 8, Appl1
13	6732	98.8	1255	US-08-484-438-8	Sequence 8, Appl1
14	3630	53.3	782	US-09-146-283-4	Sequence 4, Appl1
15	3630	53.3	782	US-08-579-823A-4	Sequence 4, Appl1
16	3630	53.3	782	US-09-344-195-4	Sequence 4, Appl1
17	3590	52.7	645	US-09-602-812A-13	Sequence 13, Appl1
18	3473	51.0	624	US-08-422-108-1	Sequence 1, Appl1
19	3473	51.0	624	US-08-422-734-1	Sequence 1, Appl1
20	3167	46.5	1210	US-08-475-035-4	Sequence 7, Appl1
21	3167	46.5	1210	US-08-484-438-7	Sequence 7, Appl1
22	3167	46.5	1210	US-09-715-249-2	Sequence 2, Appl1
23	3146	46.2	1210	US-09-723-307-67	Sequence 67, Appl1
24	3091	45.4	580	US-08-414-417B-69	Sequence 69, Appl1
25	3091	45.4	580	US-08-486-348A-69	Sequence 69, Appl1
26	3091	45.4	580	US-08-468-545B-69	Sequence 69, Appl1
27	3091	45.4	580	US-08-466-680B-69	Sequence 69, Appl1

28	3091	45.4	580	2	US-09-354-533-69	Sequence 69, Appl1
29	3004.5	44.1	1308	1	US-08-484-438-2	Sequence 2, Appl1
30	2871.5	42.1	1058	1	US-08-484-438-4	Sequence 4, Appl1
31	2441.5	35.8	1342	1	US-07-978-895-4	Sequence 4, Appl1
32	2441.5	35.8	1342	1	US-08-484-438-9	Sequence 9, Appl1
33	2441.5	35.8	1342	1	US-08-473-119-4	Sequence 4, Appl1
34	2441.5	35.8	1342	1	US-08-475-352-4	Sequence 4, Appl1
35	2441.5	35.8	1342	2	US-09-170-699-4	Sequence 4, Appl1
36	2441.5	35.8	1360	2	US-09-949-016-8022	Sequence 8022, Ap
37	2434	35.7	1343	6	5183884-4	Patent No. 5183884
38	1878	27.6	419	2	US-09-630-155-2	Sequence 2, Appl1
39	1646.5	24.2	705	1	US-08-456-647B-4	Sequence 4, Appl1
40	1646.5	24.2	705	1	US-08-237-401A-4	Sequence 4, Appl1
41	1532.5	22.5	644	1	US-08-336-708A-9	Sequence 9, Appl1
42	1451	21.3	911	1	US-08-484-438-10	Sequence 10, Appl1
43	1389.5	20.4	265	1	US-07-857-224B-66	Sequence 66, Appl1
44	1235	18.1	279	1	US-08-701-191A-14	Sequence 14, Appl1
45	1235	18.1	279	2	US-09-664-526-14	Sequence 14, Appl1

ALIGNMENTS

RESULT 1
US-08-625-101-2
; Sequence 2, Application US/08625101
; Patent No. 5859445
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
; TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
; TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/625,101
; FILING DATE: 01-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULAR TYPE: protein
; US-08-625-101-2
Query Match 100.0%; Score 6815; DB 1; Length 1255;
Best Local Similarity 100.0%; Pred. No: 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 METALCRWGLLALLPFGAASVYCTGTDMKRLRLPASPETHLDMRLHYOGCQVQGNL 60
DB 1 METALCRWGLLALLPFGAASVYCTGTDMKRLRLPASPETHLDMRLHYOGCQVQGNL 60

```

QY 61 ELTYLPTNASISFLDIOEVQGYVLIANQVROVPLQRLIRVGTOLFEDNYALAVLNG 120
DB 61 ELTYLPTNASISFLDIOEVQGYVLIANQVROVPLQRLIRVGTOLFEDNYALAVLNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSITELIKGGVLIQRNPOLCYQDTIIMKDI FHKQNOLA 180
DB 121 DPLNNTTPTVTGASPGGLRELQRLSITELIKGGVLIQRNPOLCYQDTIIMKDI FHKQNOLA 180
QY 181 LTLIDTNSRACHPSPMCKSGRCWGESSEDCQSITRTVACAGGARGCKGPLPTDCCHBQC 240
DB 181 LTLIDTNSRACHPSPMCKSGRCWGESSEDCQSITRTVACAGGARGCKGPLPTDCCHBQC 240
QY 241 AAGCTGPKHSDCLAFHFNHSGICELHCPALVTYNTDTFESNPNBGRVTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLAFHFNHSGICELHCPALVTYNTDTFESNPNBGRVTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARVYCGIGMHLREVRAVTSAN 360
DB 301 YNYLSTDVGSCTLVCPHLNQEVTAEDGTQRCCKSKPCARVYCGIGMHLREVRAVTSAN 360
QY 361 IOEPAGCKKIFGSLAFLPESFPDPAASNTAPLOPQLOVFEETLEITGYLISAMPDSL 420
DB 361 IOEPAGCKKIFGSLAFLPESFPDPAASNTAPLOPQLOVFEETLEITGYLISAMPDSL 420
QY 421 DLSVFQNLQVIRGRILHNGAVSLTLOGLISWLGRLSRLGSGALIHNNTHLCFVATV 480
DB 421 DLSVFQNLQVIRGRILHNGAVSLTLOGLISWLGRLSRLGSGALIHNNTHLCFVATV 480
QY 481 PWDOLFRRPHQALLHTANRPEDECVBEGGLACHQLCARHGCWPGTQCVCNGQFIRGQEC 540
DB 481 PWDOLFRRPHQALLHTANRPEDECVBEGGLACHQLCARHGCWPGTQCVCNGQFIRGQEC 540
QY 541 VEECVLQGLPREYVNAHRCLEPCHECOPONGSVTCFGEADQCACAHYKPPPCVARG 600
DB 541 VEECVLQGLPREYVNAHRCLEPCHECOPONGSVTCFGEADQCACAHYKPPPCVARG 600
QY 601 PSQVAPDLSYMBIMKFPDEBGAQCPINCTHSCVDLDDKGPABORASPLTISIAYVG 660
DB 601 PSQVAPDLSYMBIMKFPDEBGAQCPINCTHSCVDLDDKGPABORASPLTISIAYVG 660
QY 661 ILVVVVLGVVFGILLIKRQOKIRKTYMRLLOETELVEPLTSGAMPQAOQRILKEITEL 720
DB 661 ILVVVVLGVVFGILLIKRQOKIRKTYMRLLOETELVEPLTSGAMPQAOQRILKEITEL 720
QY 721 RKVKVLGSGAFGTWKGIWIPDGENVKIPVAIKVIRENTSPANKELIDEAVYMAGVGSP 780
DB 721 RKVKVLGSGAFGTWKGIWIPDGENVKIPVAIKVIRENTSPANKELIDEAVYMAGVGSP 780
QY 781 YVSRLLGICLTSTVQLVTLQMPYGLLDHVRNRRGLSGQDILLNMCQIAKMSYLEDVR 840
DB 781 YVSRLLGICLTSTVQLVTLQMPYGLLDHVRNRRGLSGQDILLNMCQIAKMSYLEDVR 840
QY 841 LVHRLAARNVLVKSPPNHVKTIDFGLARLLDIDETRYHADGKVPKIMMALESILRRFT 900
DB 841 LVHRLAARNVLVKSPPNHVKTIDFGLARLLDIDETRYHADGKVPKIMMALESILRRFT 900
QY 901 HOSDWASVGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYIMVKKCM 960
DB 901 HOSDWASVGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPICTIDVYIMVKKCM 960
QY 961 IDSECRPRRELVSERSMARDPQRFVVIQNEEDLPASPLDSTFYRSLLEDDMDGLVDA 1020
DB 961 IDSECRPRRELVSERSMARDPQRFVVIQNEEDLPASPLDSTFYRSLLEDDMDGLVDA 1020
QY 1021 BEYLVPOQGFPCPDPAFGAGVHHRHSSSTRSGGDLTGLLEPSEBAPRSPAPBSG 1080
DB 1021 BEYLVPOQGFPCPDPAFGAGVHHRHSSSTRSGGDLTGLLEPSEBAPRSPAPBSG 1080
QY 1081 AGSDVFDGDLGMAKAGLSLPTHPDSPLORYSESDPTVLPSETGTGYVAPLTCSPQPEV 1140
DB 1081 AGSDVFDGDLGMAKAGLSLPTHPDSPLORYSESDPTVLPSETGTGYVAPLTCSPQPEV 1140

```

```

QY 1141 NOPVRRPQPSBREGPLPAAPAGATLERPKTSLSPKNGVYKDVAFAGAVENPEYLTPO 1200
DB 1141 NOPVRRPQPSBREGPLPAAPAGATLERPKTSLSPKNGVYKDVAFAGAVENPEYLTPO 1200
QY 1201 GGAAPQHPAPFAPFADNLYYWDODPERGAPSTFGPTAENPEYLGIDVPV 1255
DB 1201 GGAAPQHPAPFAPFADNLYYWDODPERGAPSTFGPTAENPEYLGIDVPV 1255

RESULT 2
US-08-356-786-2
; Sequence 2, Application US/08356786
; Patent No. 587305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pletcher, Testra, Hurwitz, & Thibeault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

;
; SOFTWARE: Patent in release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pletcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; TELEPHONE/DOCKET NUMBER: CRP-053
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-786-2

Query Match 100.0%; Score 6815; DB 1; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 181 LTLIDTNRBRACHPSPCKSGSRGWGSESDCOSLTRITVCAGGACARCKGPLETDCCHBQC 240
Qy 241 AAGCTGPRGSDCLACLRHNSGICELHCPALVTYNTDTPESMPNBEGRYTPASCVTACP 300
Db 241 AAGCTGPRGSDCLACLRHNSGICELHCPALVTYNTDTPESMPNBEGRYTPASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPLNQOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVRVAVTSAN 360
Db 301 YNYLSTDVGSCTLVCPLNQOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVRVAVTSAN 360
Qy 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLOPBOLOVFETLEBITGYLYISAMPDSL 420
Db 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLOPBOLOVFETLEBITGYLYISAMPDSL 420
Qy 421 DLSVFQNTQVTRGRILHNGAYSLTLOGLGISWLGRLSRLBELSGSLIHHNTHLCFVHTV 480
Db 421 DLSVFQNTQVTRGRILHNGAYSLTLOGLGISWLGRLSRLBELSGSLIHHNTHLCFVHTV 480
Qy 481 PMDOLFRNPHOALHTTANRPEDECVGEGLAGHQLCARGHCWGPPTQCVCNSQFLRGQEC 540
Db 481 PMDOLFRNPHOALHTTANRPEDECVGEGLAGHQLCARGHCWGPPTQCVCNSQFLRGQEC 540
Qy 541 VEBECVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFCVANC 600
Db 541 VEBECVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFCVANC 600
Qy 601 PSQVPRDLSYMPIMKFPDEBGAQCPINCHTSCVDLDKGPAPORASPLTSIIISAVNG 660
Db 601 PSQVPRDLSYMPIMKFPDEBGAQCPINCHTSCVDLDKGPAPORASPLTSIIISAVNG 660
Qy 661 ILLVVVLGVFGIILKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOAQRILIKETEL 720
Db 661 ILLVVVLGVFGIILKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOAQRILIKETEL 720
Qy 721 RKVKVLGSGAFGTYYKGIWIPDGENVKIPVALKVLRENTSPKANEIIDEAYVMAVGSP 780
Db 721 RKVKVLGSGAFGTYYKGIWIPDGENVKIPVALKVLRENTSPKANEIIDEAYVMAVGSP 780
Qy 781 YVSRLLGICLTSTVOLVQLMFYGCLLDHVRNENRGLSQDILLNMCQIAKMSYLEBVR 840
Db 781 YVSRLLGICLTSTVOLVQLMFYGCLLDHVRNENRGLSQDILLNMCQIAKMSYLEBVR 840
Qy 841 LVHRLLARNVLYKSPNHVKITDPGLARLLDIDETEYHADGGKVPILKMALESILRRPT 900
Db 841 LVHRLLARNVLYKSPNHVKITDPGLARLLDIDETEYHADGGKVPILKMALESILRRPT 900
Qy 901 HOSDVMSYGVTVWELMTFGAKPYDGI PAEIPDLLEKGERLPOPPICITIDVYMIWVKCM 960
Db 901 HOSDVMSYGVTVWELMTFGAKPYDGI PAEIPDLLEKGERLPOPPICITIDVYMIWVKCM 960
Qy 961 IDSECRPRPRELVSSESRMARDPORFVYIQNEDLGPASPLDSTFYRSLLEDDMDGLDYA 1020
Db 961 IDSECRPRPRELVSSESRMARDPORFVYIQNEDLGPASPLDSTFYRSLLEDDMDGLDYA 1020
Qy 1021 BEVLYPQOGFCPPDPAFGAGVHHRRHSSSTRSGGDLTLGLBPSEBEPSPPLAPSBG 1080
Db 1021 BEVLYPQOGFCPPDPAFGAGVHHRRHSSSTRSGGDLTLGLBPSEBEPSPPLAPSBG 1080
Qy 1081 AGSDVFDGDLGWAAGLQSLPTHDPSPLQRYSEDPVYLPSETDGYVAPLTCSPQPEVY 1140
Db 1081 AGSDVFDGDLGWAAGLQSLPTHDPSPLQRYSEDPVYLPSETDGYVAPLTCSPQPEVY 1140
Qy 1141 NOPDVRPOPSPREGRPLPAARPAAGATLEPRTLSPKNGVAVDVAFGAVENPEYLTPQ 1200
Db 1141 NOPDVRPOPSPREGRPLPAARPAAGATLEPRTLSPKNGVAVDVAFGAVENPEYLTPQ 1200
Qy 1201 GGAAPOPHPPAFSPADNLYYMDODPBERGAPSTFKGTPTAENPEYLGLOVPV 1255
Db 1201 GGAAPOPHPPAFSPADNLYYMDODPBERGAPSTFKGTPTAENPEYLGLOVPV 1255

```

RESULT 3
US-09-441-411-6

```

; Sequence 6, Application US/09441411
; Patent No. 6734172
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary U.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441,411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-411-6

Query Match      100.0%; Score 6815; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No: 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MELAALCRWGLLALLPFGAASVQVCTGTDKRLIPASBETHLDMRLHYGCGVVGNTL 60
Db 1 MELAALCRWGLLALLPFGAASVQVCTGTDKRLIPASBETHLDMRLHYGCGVVGNTL 60
Qy 61 ETTYPTNASLFLDIDIOEVGYVLIANNQVQVLOLRIYRQVLPEDNVALAVLDNG 120
Db 61 ETTYPTNASLFLDIDIOEVGYVLIANNQVQVLOLRIYRQVLPEDNVALAVLDNG 120
Qy 121 DPLANTTPTVAGSPGLRELOLRSLTEILKGSVLIOBNPOLCYOPTILMKDIFHKNQOLA 180
Db 121 DPLANTTPTVAGSPGLRELOLRSLTEILKGSVLIOBNPOLCYOPTILMKDIFHKNQOLA 180
Qy 181 LTLIDTNRBRACHPSPCKSGSRGWGSESDCOSLTRITVCAGGACARCKGPLETDCCHBQC 240
Db 181 LTLIDTNRBRACHPSPCKSGSRGWGSESDCOSLTRITVCAGGACARCKGPLETDCCHBQC 240
Qy 241 AAGCTGPRGSDCLACLRHNSGICELHCPALVTYNTDTPESMPNBEGRYTPASCVTACP 300
Db 241 AAGCTGPRGSDCLACLRHNSGICELHCPALVTYNTDTPESMPNBEGRYTPASCVTACP 300
Qy 301 YNYLSTDVGSCTLVCPLNQOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVRVAVTSAN 360
Db 301 YNYLSTDVGSCTLVCPLNQOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVRVAVTSAN 360
Qy 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLOPBOLOVFETLEBITGYLYISAMPDSL 420
Db 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLOPBOLOVFETLEBITGYLYISAMPDSL 420
Qy 421 DLSVFQNTQVTRGRILHNGAYSLTLOGLGISWLGRLSRLBELSGSLIHHNTHLCFVHTV 480
Db 421 DLSVFQNTQVTRGRILHNGAYSLTLOGLGISWLGRLSRLBELSGSLIHHNTHLCFVHTV 480
Qy 481 PMDOLFRNPHOALHTTANRPEDECVGEGLAGHQLCARGHCWGPPTQCVCNSQFLRGQEC 540
Db 481 PMDOLFRNPHOALHTTANRPEDECVGEGLAGHQLCARGHCWGPPTQCVCNSQFLRGQEC 540
Qy 541 VEBECVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFCVANC 600
Db 541 VEBECVLOGLPREYVNAHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFCVANC 600
Qy 601 PSQVPRDLSYMPIMKFPDEBGAQCPINCHTSCVDLDKGPAPORASPLTSIIISAVNG 660
Db 601 PSQVPRDLSYMPIMKFPDEBGAQCPINCHTSCVDLDKGPAPORASPLTSIIISAVNG 660
Qy 661 ILLVVVLGVFGIILKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOAQRILIKETEL 720
Db 661 ILLVVVLGVFGIILKRRQOKIRKTYMRLLQETELVEPLTPSGAMPNOAQRILIKETEL 720
Qy 721 RKVKVLGSGAFGTYYKGIWIPDGENVKIPVALKVLRENTSPKANEIIDEAYVMAVGSP 780

```

```

Db      721 RKVKVLSGSAFGTYVYKGIWIPDGENVKIPVAIKVIRENTSPKANKEILDEAVYVAGVSP 780
Qy      781 YVSRLLGICLTSTVQVLTQMLPVGCLLDHVENRGRLSQDLNMCQIAGMSYLEBYR 840
Db      781 YVSRLLGICLTSTVQVLTQMLPVGCLLDHVENRGRLSQDLNMCQIAGMSYLEBYR 840
Qy      841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETEHADGKVPDKMMALLESILRRFT 900
Db      841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETEHADGKVPDKMMALLESILRRFT 900
Qy      901 HQSDVWSGYVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITIDVYMIWVKCM 960
Db      901 HQSDVWSGYVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITIDVYMIWVKCM 960
Qy      961 IDSECRPFRELVSFESRMARDPQRFVYIQNEDLPASPLDSTFYRSLLDDMDKDLVDA 1020
Db      961 IDSECRPFRELVSFESRMARDPQRFVYIQNEDLPASPLDSTFYRSLLDDMDKDLVDA 1020
Qy      1021 EBYLVPQGGFCPPDAPAGAGVHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
Db      1021 EBYLVPQGGFCPPDAPAGAGVHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSEG 1080
Qy      1081 AGSDVPDDDLGMAKAGIQSLPTHDPSPLOQYSBDPTVPLBSETDGYVAPLTCSPQPEYV 1140
Db      1081 AGSDVPDDDLGMAKAGIQSLPTHDPSPLOQYSBDPTVPLBSETDGYVAPLTCSPQPEYV 1140
Qy      1141 NOPDVRPOPSPREBPLPAAPAGATLEBPKTLSPGKGVVQDVAFGAGAVENPEYLTPO 1200
Db      1141 NOPDVRPOPSPREBPLPAAPAGATLEBPKTLSPGKGVVQDVAFGAGAVENPEYLTPO 1200
Qy      1201 GGAAPQHPPPAPFPAFNLVYWDODPPBPGAPBSTFGKTPAENPEYLGIDVY 1255
Db      1201 GGAAPQHPPPAPFPAFNLVYWDODPPBPGAPBSTFGKTPAENPEYLGIDVY 1255

```

RESULT 4

US-09-167-516-2
Sequence 2, Application US/09167516
Patent No. 6953573

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.
APPLICANT: Disig, Mary L.
TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/167,516
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/625,101
FILING DATE: 01-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-167-516-2

Query Match 100.0%; Score 6815; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MELAALCRWGLILALLPPGAASVQCTGTMDKRLPASPETHLMDLRHYGCGVQGNL 60
Db      1 MELAALCRWGLILALLPPGAASVQCTGTMDKRLPASPETHLMDLRHYGCGVQGNL 60
Qy      61 ELTYLPTNASLSTFQDIOEVGVGVLIANNOVRQVPLQRLIVRGITQLFEDNYALAVLDNG 120
Db      61 ELTYLPTNASLSTFQDIOEVGVGVLIANNOVRQVPLQRLIVRGITQLFEDNYALAVLDNG 120
Qy      121 DPLNNTTPVTGASPGGLREIQLRLSTEILKGVLIQRNPOLCYODTILMKDI FPKNNQLA 180
Db      121 DPLNNTTPVTGASPGGLREIQLRLSTEILKGVLIQRNPOLCYODTILMKDI FPKNNQLA 180
Qy      181 LTLIDTNRSPACHPSCSPCKGSRGWSSESDQSLTRTYCAGGCAKCGPLPTDCHEQC 240
Db      181 LTLIDTNRSPACHPSCSPCKGSRGWSSESDQSLTRTYCAGGCAKCGPLPTDCHEQC 240
Qy      241 AAGCTGPRHSQCLACLPFNHSGICELHCPALVTYNTDFFESMPBEGRYTGASCVTACP 300
Db      241 AAGCTGPRHSQCLACLPFNHSGICELHCPALVTYNTDFFESMPBEGRYTGASCVTACP 300
Qy      301 YNVLSTDVSGCTLYVCPHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
Db      301 YNVLSTDVSGCTLYVCPHNOEVTAEADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
Qy      361 IQEFAGCKKIFGSLAFIPESFDGPPASNTAPLOEPOLOVEFTLEITGYLYISMPDLSL 420
Db      361 IQEFAGCKKIFGSLAFIPESFDGPPASNTAPLOEPOLOVEFTLEITGYLYISMPDLSL 420
Qy      421 DLSVFONLQVIRGRILNNGASLTLOGIGIFWLGIRSLREISGGLALIHNTHTCFYTV 480
Db      421 DLSVFONLQVIRGRILNNGASLTLOGIGIFWLGIRSLREISGGLALIHNTHTCFYTV 480
Qy      481 PMDQLFNPHQALHTANRDECEVGBGLACHOLACARHCWGPPTQVNCQSQFLRGEC 540
Db      481 PMDQLFNPHQALHTANRDECEVGBGLACHOLACARHCWGPPTQVNCQSQFLRGEC 540
Qy      541 VEECRVLOGLPREVYVNAHCLPCHPECOPONGSVTCGPEADQVACAHYDPPFCVARC 600
Db      541 VEECRVLOGLPREVYVNAHCLPCHPECOPONGSVTCGPEADQVACAHYDPPFCVARC 600
Qy      601 PSQVAPDLSTYMPKFPDEBGACOPCPINCHSCVDLDDKCPAPBORASPLTSIIISAVG 660
Db      601 PSQVAPDLSTYMPKFPDEBGACOPCPINCHSCVDLDDKCPAPBORASPLTSIIISAVG 660
Qy      661 ILVVVLGVVVGILIKRQOKIRKTYMKRLLQETLEVEPLTPSGAMPQOAMRLIKETEL 720
Db      661 ILVVVLGVVVGILIKRQOKIRKTYMKRLLQETLEVEPLTPSGAMPQOAMRLIKETEL 720
Qy      721 RKVKVLSGSAFGTYVYKGIWIPDGENVKIPVAIKVIRENTSPKANKEILDEAVYVAGVSP 780
Db      721 RKVKVLSGSAFGTYVYKGIWIPDGENVKIPVAIKVIRENTSPKANKEILDEAVYVAGVSP 780
Qy      781 YVSRLLGICLTSTVQVLTQMLPVGCLLDHVENRGRLSQDLNMCQIAGMSYLEBYR 840
Db      781 YVSRLLGICLTSTVQVLTQMLPVGCLLDHVENRGRLSQDLNMCQIAGMSYLEBYR 840
Qy      841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETEHADGKVPDKMMALLESILRRFT 900
Db      841 LVHRDLAARNVLYKSPNHVKITDFGLARLLDIDETEHADGKVPDKMMALLESILRRFT 900
Qy      901 HQSDVWSGYVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITIDVYMIWVKCM 960

```


Db 901 HQSDWSGYVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITIDVYIMVKCM 960
 Qy 961 IDSECRPRFRELVSFSEFMAADPQRFVYIQNEDLGAPSLDSTFYRSLLEDDMGDLVDA 1020
 Db 961 IDSECRPRFRELVSFSEFMAADPQRFVYIQNEDLGAPSLDSTFYRSLLEDDMGDLVDA 1020
 Qy 1021 BEYLVPOQGFPCPDPAAGAGMHHRRSSSTRSGGGLTLGLEPSEEARPSPLAPSEG 1080
 Db 1021 BEYLVPOQGFPCPDPAAGAGMHHRRSSSTRSGGGLTLGLEPSEEARPSPLAPSEG 1080
 Qy 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYV 1140
 Db 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYV 1140
 Qy 1141 NOPDVRPOPSPRPGPLPAAPAGATLERPTLSPGKGVYKDVAFAGAVENPEYLTPO 1200
 Db 1141 NOPDVRPOPSPRPGPLPAAPAGATLERPTLSPGKGVYKDVAFAGAVENPEYLTPO 1200
 Qy 1201 GGAAPQHPAPFSPAPFNDLYYMODDPPERGAPSTFGTPTAENPEYLGLDVPE 1255
 Db 1201 GGAAPQHPAPFSPAPFNDLYYMODDPPERGAPSTFGTPTAENPEYLGLDVPE 1255

RESULT 5

US-08-467-083-68
 Sequence 68, Application US/08467083
 Patent No. 5726023
 GENERAL INFORMATION:
 APPLICANT: Cheever, Martin A.
 APPLICANT: Disla, Mary L.
 TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
 TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
 NUMBER OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Seed and Berry
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: US
 ZIP: 98104-7092
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,083
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/414,417
 FILING DATE: 06-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Sharkey, Richard G.
 REGISTRATION NUMBER: 32,629
 REFERENCE/DOCKET NUMBER: 920010.448C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 TELEX: 3723836 SEEDANBERY
 INFORMATION FOR SEQ ID NO: 68:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1255 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-467-083-68

Query Match 99.9%; Score 6806; DB 1; Length 1255;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MELAALCRWGLLALLPPGAASVCTGTCTGDKRLRPASPEITHLMDLRLLYQSGCVQGNL 60
 Db 1 MELAALCRWGLLALLPPGAASVCTGTCTGDKRLRPASPEITHLMDLRLLYQSGCVQGNL 60
 Qy 61 ELTYLPTNASLSPFQDIOEVQGYVLIANQVROYPLQRLIRVGTQLFEDNYALAVLNG 120
 Db 61 ELTYLPTNASLSPFQDIOEVQGYVLIANQVROYPLQRLIRVGTQLFEDNYALAVLNG 120
 Qy 121 DPLNNTTPYTGASGGGLREIQLRSLTEILKGGVLIQRNPOLCYODTILMKDI FHKNNOLA 180
 Db 121 DPLNNTTPYTGASGGGLREIQLRSLTEILKGGVLIQRNPOLCYODTILMKDI FHKNNOLA 180
 Qy 181 LTLIDTNRSRACHSPWCCKSRCKWSSSEDCSLTRTVACGACARCKGPLPTDCHEQC 240
 Db 181 LTLIDTNRSRACHSPWCCKSRCKWSSSEDCSLTRTVACGACARCKGPLPTDCHEQC 240
 Qy 241 AAGCTGPHSDCLALPHNHSICELHCPALVTYNTDTFESMPNBEGRYTTGASCVTACP 300
 Db 241 AAGCTGPHSDCLALPHNHSICELHCPALVTYNTDTFESMPNBEGRYTTGASCVTACP 300
 Qy 301 YNYLSTDVGSCTIACPLHNOEVTAEADGTQCEKCSKPCARVCYGLGMEHLEVRVAVTSAN 360
 Db 301 YNYLSTDVGSCTIACPLHNOEVTAEADGTQCEKCSKPCARVCYGLGMEHLEVRVAVTSAN 360
 Qy 361 IQEPPAGCKIFGSLAFLPESFPDGPASNTAPLOEOLQVFETLEITGYLYISAMPDPLP 420
 Db 361 IQEPPAGCKIFGSLAFLPESFPDGPASNTAPLOEOLQVFETLEITGYLYISAMPDPLP 420
 Qy 421 DLSVFQNLQVIRGILNNGAVSLTLOGLISWLGRLSIRELSGLAIHNHTHLCFVHTV 480
 Db 421 DLSVFQNLQVIRGILNNGAVSLTLOGLISWLGRLSIRELSGLAIHNHTHLCFVHTV 480
 Qy 481 PMDQLFRRPHQALHTNRPDEDCVGEGLACHQICARHCMPGPTQCVNCSQFLRGEC 540
 Db 481 PMDQLFRRPHQALHTNRPDEDCVGEGLACHQICARHCMPGPTQCVNCSQFLRGEC 540
 Qy 541 VEECRVLQGLREYVNAHCLPCHPECOPOKGSVTCFEPZADQVACAHYDPPFCVARC 600
 Db 541 VEECRVLQGLREYVNAHCLPCHPECOPOKGSVTCFEPZADQVACAHYDPPFCVARC 600
 Qy 601 PSQVYKPDLSYMPIMKFPDEBGACQPCPINTCHSCVDLDDKCPABORASPLTISAVYG 660
 Db 601 PSQVYKPDLSYMPIMKFPDEBGACQPCPINTCHSCVDLDDKCPABORASPLTISAVYG 660
 Qy 661 ILLVYVIGVFGIILIKRQOKIRKCYTMRRLQETELVEPLTPSGAMPQAOAMRLIKETEL 720
 Db 661 ILLVYVIGVFGIILIKRQOKIRKCYTMRRLQETELVEPLTPSGAMPQAOAMRLIKETEL 720
 Qy 721 RKVAVLGSAGFYVYKGIWIPDGENVKIPVAIKYLRRENTSPKANKELIDEAYVVAAGVSP 780
 Db 721 RKVAVLGSAGFYVYKGIWIPDGENVKIPVAIKYLRRENTSPKANKELIDEAYVVAAGVSP 780
 Qy 781 YVSRILGICLSTVQVLTQMPYGCCLLDHVENRGLSGODLNMQIAKMSYLEBVR 840
 Db 781 YVSRILGICLSTVQVLTQMPYGCCLLDHVENRGLSGODLNMQIAKMSYLEBVR 840
 Qy 841 LVHRDLAARNLVYVSPNHYKTTDGLARLLDIDETVYHADGSKPIKMALESLIRRF 900
 Db 841 LVHRDLAARNLVYVSPNHYKTTDGLARLLDIDETVYHADGSKPIKMALESLIRRF 900
 Qy 901 HQSDWSGYVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITIDVYIMVKCM 960
 Db 901 HQSDWSGYVTWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITIDVYIMVKCM 960
 Qy 961 IDSECRPRFRELVSFSEFMAADPQRFVYIQNEDLGAPSLDSTFYRSLLEDDMGDLVDA 1020
 Db 961 IDSECRPRFRELVSFSEFMAADPQRFVYIQNEDLGAPSLDSTFYRSLLEDDMGDLVDA 1020
 Qy 1021 BEYLVPOQGFPCPDPAAGAGMHHRRSSSTRSGGGLTLGLEPSEEARPSPLAPSEG 1080
 Db 1021 BEYLVPOQGFPCPDPAAGAGMHHRRSSSTRSGGGLTLGLEPSEEARPSPLAPSEG 1080
 Qy 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYV 1140

|||||
Db 1081 AGSVDFGDLGMGAAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQBEYV 1140
Qy 1141 NOPVRRPOPSPRRSGPLPAARPAAGATLERPKTSLSGKGVYKDVAFAGGAVENBEYLTPO 1200
Db 1141 NOPVRRPOPSPRRSGPLPAARPAAGATLERPKTSLSGKGVYKDVAFAGGAVENBEYLTPO 1200
Qy 1201 GAAPQHPHPPAPFAPFNDLYYWDODPPERGAPSTFGKPTAENPEYLGIDVYV 1255
Db 1201 GAAPQHPHPPAPFAPFNDLYYWDODPPERGAPSTFGKPTAENPEYLGIDVYV 1255

RESULT 6
US-08-414-117B-68
Sequence 68, Application US/08414417B
Patent No. 5801005
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disla, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/414,417B
FILING DATE: 31-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4500
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-414-117B-68

Query Match 99.9%; Score 6806; DB 1; Length 1255;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MELAALCRWGLLLALLPPGASTOVCTGDMKRLRPASPEETHLMDLRHLVYGCQVVGNTL 60
Db 1 MELAALCRWGLLLALLPPGASTOVCTGDMKRLRPASPEETHLMDLRHLVYGCQVVGNTL 60
Qy 61 ELTYLPTNASISFLQDIOEVGVYVLIANQVRQVPLQRLRIVRSTQVLFEDNYALAVLDNG 120
Db 61 ELTYLPTNASISFLQDIOEVGVYVLIANQVRQVPLQRLRIVRSTQVLFEDNYALAVLDNG 120
Qy 121 DPLANTTPVVGASGGGLRELQRLSLTELKGVLIQRNPOLCYQDITLMDIIFKNNOLA 180
Db 121 DPLANTTPVVGASGGGLRELQRLSLTELKGVLIQRNPOLCYQDITLMDIIFKNNOLA 180
Qy 181 LTLIDTNRSPRACHCSPWCKSGRCMGSSESDCQSLTRTVCGAGCARCKGPLPTDCCHEQC 240
Db 181 LTLIDTNRSPRACHCSPWCKSGRCMGSSESDCQSLTRTVCGAGCARCKGPLPTDCCHEQC 240

Qy 241 AAGTGRHSDCLACLFHNSHICEIHLCPALVTYNTDTFESMNPBECRYTFGASCVTACP 300
Db 241 AAGTGRHSDCLACLFHNSHSGICEIHLCPALVTYNTDTFESMNPBECRYTFGASCVTACP 300
Qy 301 YNTISTVGSCTTVCPLAHNOEVTAEEDTORCEKSKCAVVCVGLGMEHLREVAVLVSAN 360
Db 301 YNTISTVGSCTTVCPLAHNOEVTAEEDTORCEKSKCAVVCVGLGMEHLREVAVLVSAN 360
Qy 361 IOEFAGCKKIFGSLAFPLPESFDDGPASNTAPLQEQLOVEETELEITGYLYISAMPDLP 420
Db 361 IOEFAGCKKIFGSLAFPLPESFDDGPASNTAPLQEQLOVEETELEITGYLYISAMPDLP 420
Qy 421 DLSVFQMLQVIRGRILHNGAYSLTLQGLGISWGLSLRSLRGLALIHNTHLCPVHTV 480
Db 421 DLSVFQMLQVIRGRILHNGAYSLTLQGLGISWGLSLRSLRGLALIHNTHLCPVHTV 480
Qy 481 PMDOLFENPHQALLHTANRPEDECVGGLACHOLCANGHCMGPPTOCVNCOSQPLRGQEC 540
Db 481 PMDOLFENPHQALLHTANRPEDECVGGLACHOLCANGHCMGPPTOCVNCOSQPLRGQEC 540
Qy 541 VEEGRVLOGLPREYVNAHGLPCHPECOPONGSVTCFGEPAADQVACAHYKDPFCVARC 600
Db 541 VEEGRVLOGLPREYVNAHGLPCHPECOPONGSVTCFGEPAADQVACAHYKDPFCVARC 600
Qy 601 PSGVKPDLSTMPYIWKFPDEGACQPCPINCTHSCVDLDDYGCAPABOPASPLTISAVVG 660
Db 601 PSGVKPDLSTMPYIWKFPDEGACQPCPINCTHSCVDLDDYGCAPABOPASPLTISAVVG 660
Qy 661 ILVYVTVGVVFGILIRKROOKIRKYMRLLOETELVEPLTPSGAMNOAMRLKTEL 720
Db 661 ILVYVTVGVVFGILIRKROOKIRKYMRLLOETELVEPLTPSGAMNOAMRLKTEL 720
Qy 721 RKVKVLSGAGFVYKGIWIIPDEBNVXI PVAIKVLRNTPSPKANKELIDEAYVAVAGVSP 780
Db 721 RKVKVLSGAGFVYKGIWIIPDEBNVXI PVAIKVLRNTPSPKANKELIDEAYVAVAGVSP 780
Qy 781 YVSRBLGICLTSTYQVLTQMPYGCCLDHYRENRGRSGODLNMWCMQIAKMSYLEYR 840
Db 781 YVSRBLGICLTSTYQVLTQMPYGCCLDHYRENRGRSGODLNMWCMQIAKMSYLEYR 840
Qy 841 LVHRDLAARVNLVKSPPNHVXITDGLARLIDIBETEVHAGGVKPIKMALESTLRRFT 900
Db 841 LVHRDLAARVNLVKSPPNHVXITDGLARLIDIBETEVHAGGVKPIKMALESTLRRFT 900
Qy 901 HQSDVMSYGVTVLWELMTFGAKPYDGI PAREIPDLLEKGERLPORPCTIDVYMIWVCM 960
Db 901 HQSDVMSYGVTVLWELMTFGAKPYDGI PAREIPDLLEKGERLPORPCTIDVYMIWVCM 960
Qy 961 IDSECRPRFRELVSERSMARDPQRFVYIQNEDELGPASPLDSTFYRSLLEDDMDGLVDA 1020
Db 961 IDSECRPRFRELVSERSMARDPQRFVYIQNEDELGPASPLDSTFYRSLLEDDMDGLVDA 1020
Qy 1021 EEYLVPOQGFPCDPPAPGAGGMVHHRRSSSTRSGGDLTLGLEPSEERA PRSPLAPSEG 1080
Db 1021 EEYLVPOQGFPCDPPAPGAGGMVHHRRSSSTRSGGDLTLGLEPSEERA PRSPLAPSEG 1080
Qy 1081 AGSDVDFGDLGMGAAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQBEYV 1140
Db 1081 AGSDVDFGDLGMGAAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQBEYV 1140
Qy 1141 NOPVRRPOPSPRRSGPLPAARPAAGATLERPKTSLSGKGVYKDVAFAGGAVENBEYLTPO 1200
Db 1141 NOPVRRPOPSPRRSGPLPAARPAAGATLERPKTSLSGKGVYKDVAFAGGAVENBEYLTPO 1200
Qy 1201 GAAPQHPHPPAPFAPFNDLYYWDODPPERGAPSTFGKPTAENPEYLGIDVYV 1255
Db 1201 GAAPQHPHPPAPFAPFNDLYYWDODPPERGAPSTFGKPTAENPEYLGIDVYV 1255

RESULT 7
US-08-486-348A-68
Sequence 68, Application US/08486348A
Patent No. 5846538

```

GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,348A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-348A-68

Query Match 99.9%; Score 6806; DB 1; Length 1255;
Beet Local Similarity 99.9%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

1 MELALCMWGLLALPPGAASVQCTGDMRLPASPETHLDMLRLYQSCQVQGNL 60
1 MELALCMWGLLALPPGAASVQCTGDMRLPASPETHLDMLRLYQSCQVQGNL 60
61 ELTYLPTNASLFLQDIOEVQGYVLIANQVQVPLQRLIVRGTOLEFDNYALAVLNG 120
61 ELTYLPTNASLFLQDIOEVQGYVLIANQVQVPLQRLIVRGTOLEFDNYALAVLNG 120
121 DPLNNTTGTGASPGGLRELQRLSLTEILKGVLIQRNPQLCYQDTILMKDIFHANNOLA 180
121 DPLNNTTGTGASPGGLRELQRLSLTEILKGVLIQRNPQLCYQDTILMKDIFHANNOLA 180
181 LTLIDTNSRACHPCSPCKSGRCWGESSEDCQSLTRTYCAGGACARCKGPLETDCCHEQC 240
181 LTLIDTNSRACHPCSPCKSGRCWGESSEDCQSLTRTYCAGGACARCKGPLETDCCHEQC 240
241 AAGCTGPRHSDCLACIHRNHSICELACPALVTYNTDPESMPNDEGRYTFGASCVTACP 300
241 AAGCTGPRHSDCLACIHRNHSICELACPALVTYNTDPESMPNDEGRYTFGASCVTACP 300
301 YNYLSTDVSGCTLVGPLNNOEVTADGTORCKCKSPCARVCGYGLGMEHLREVRVATSAN 360
301 YNYLSTDVSGCTLVGPLNNOEVTADGTORCKCKSPCARVCGYGLGMEHLREVRVATSAN 360
361 IQEPAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVFETLEBITGYLISAMPDILP 420
361 IQEPAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVFETLEBITGYLISAMPDILP 420
421 DLSVFQNTQVIRGRILHNGAYSLTLQGLISWLGSLRLRELQSGIALIHNTHLCFVHTV 480
421 DLSVFQNTQVIRGRILHNGAYSLTLQGLISWLGSLRLRELQSGIALIHNTHLCFVHTV 480

```

```

481 PWDQFRNPQALHTANRPDECEVGEGLACHOLCARGHCWGPPTQCVCNSQFLRGEC 540
481 PWDQFRNPQALHTANRPDECEVGEGLACHOLCARGHCWGPPTQCVCNSQFLRGEC 540
541 VEECKVLOGLPREYNARHCLPCHECOPQNSVTCFEPBADQCACAHYDPPPCVARC 600
541 VEECKVLOGLPREYNARHCLPCHECOPQNSVTCFEPBADQCACAHYDPPPCVARC 600
601 PSQVDPDLSYMPIMKFPDEBGACOPCINCHSCVDLDDKCPARQASPLTISAYVG 660
601 PSQVDPDLSYMPIMKFPDEBGACOPCINCHSCVDLDDKCPARQASPLTISAYVG 660
661 ILVVVLGVFGLIKRRQOKIRKXTMRLLQETELVEPLTPSGAMPQAOQRIKTEL 720
661 ILVVVLGVFGLIKRRQOKIRKXTMRLLQETELVEPLTPSGAMPQAOQRIKTEL 720
721 RKVYLSGAFGTYYKGIWIPDGENVKIPVAIKVLRNTSRKANKELIDEAYVAGVSP 780
721 RKVYLSGAFGTYYKGIWIPDGENVKIPVAIKVLRNTSRKANKELIDEAYVAGVSP 780
781 YVSRLLGICTSTVQVLTQMLPYGCLDHYVENRGRLSODLLMCMQIAGMSYLBVR 840
781 YVSRLLGICTSTVQVLTQMLPYGCLDHYVENRGRLSODLLMCMQIAGMSYLBVR 840
841 LVHRDLAARNVLYKSPNHVKTITDFGLARLDIDETRYHADGAKYPIKMMALESIARRFT 900
841 LVHRDLAARNVLYKSPNHVKTITDFGLARLDIDETRYHADGAKYPIKMMALESIARRFT 900
901 HOSDWSYGVTVWEIMTFGAKPYDGI PARBIPDLLEKBERLPQPICTIDVYMIWCKM 960
901 HOSDWSYGVTVWEIMTFGAKPYDGI PARBIPDLLEKBERLPQPICTIDVYMIWCKM 960
961 IDSECRPRFRELVSFESMARDPORFVYIQNBDELSPASPLSTFRSLLEDMDMDLVA 1020
961 IDSECRPRFRELVSFESMARDPORFVYIQNBDELSPASPLSTFRSLLEDMDMDLVA 1020
1021 EBYLVPOQFCPPAPAGAGVHHRHSSTSRSGGDLTLGLBSEBEPASPLAPSEG 1080
1021 EBYLVPOQFCPPAPAGAGVHHRHSSTSRSGGDLTLGLBSEBEPASPLAPSEG 1080
1081 AGSDVFDGDLGMAKGIQSLPTHDPSPLQRYSEDPYPLPSETDGYAPLTCSQPEYV 1140
1081 AGSDVFDGDLGMAKGIQSLPTHDPSPLQRYSEDPYPLPSETDGYAPLTCSQPEYV 1140
1141 NOPDVRPQPSPREGRLPAABDAGATLERPTLSGKGVKVDVAFGAVENPEYLTPO 1200
1141 NOPDVRPQPSPREGRLPAABDAGATLERPTLSGKGVKVDVAFGAVENPEYLTPO 1200
1201 GGAAPQPPPAFSPAFDNLVYMDODPPERKAPSTFGCTPTAENPEYLGIDVPV 1255
1201 GGAAPQPPPAFSPAFDNLVYMDODPPERKAPSTFGCTPTAENPEYLGIDVPV 1255

RESULT 8
US-08-486-545B-68
Sequence 68, Application US/08468545B
Patent No. 5876712
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,545B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-466-545B-68

```

Query Match 99.9%; Score 6806; DB 1; Length 1255;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;

Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

; 1 MELAALCRWGLIALLPGAASVQCTGDMKRLPASPEHLDMLRLHYOGCQVVGQNL 60
; 1 MELAALCRWGLIALLPGAASVQCTGDMKRLPASPEHLDMLRLHYOGCQVVGQNL 60
; 61 ELTYIPINASIFLDDIOBVQGYVLIANOVROVPLORLRYRGQLFEDNYALAVLDNG 120
; 61 ELTYIPINASIFLDDIOBVQGYVLIANOVROVPLORLRYRGQLFEDNYALAVLDNG 120
; 61 ELTYIPINASIFLDDIOBVQGYVLIANOVROVPLORLRYRGQLFEDNYALAVLDNG 120
; 121 DPLNNTVTGASPGSLRELQRLTEILKGGVLIQRNPOLCYOPTIIMKDIFFKNQOLA 180
; 121 DPLNNTVTGASPGSLRELQRLTEILKGGVLIQRNPOLCYOPTIIMKDIFFKNQOLA 180
; 121 DPLNNTVTGASPGSLRELQRLTEILKGGVLIQRNPOLCYOPTIIMKDIFFKNQOLA 180
; 181 LTLIDTNRSRAHPSCPMCKSGRCWESSEDCQSLTRTVCAAGCARGCKGPLPTDCHEQC 240
; 181 LTLIDTNRSRAHPSCPMCKSGRCWESSEDCQSLTRTVCAAGCARGCKGPLPTDCHEQC 240
; 181 LTLIDTNRSRAHPSCPMCKSGRCWESSEDCQSLTRTVCAAGCARGCKGPLPTDCHEQC 240
; 241 AAGCTGPRHSDCLALHFNHSGICELHCPALVYNTDTFESNPNRSGYTGASCVTACP 300
; 241 AAGCTGPRHSDCLALHFNHSGICELHCPALVYNTDTFESNPNRSGYTGASCVTACP 300
; 241 AAGCTGPRHSDCLALHFNHSGICELHCPALVYNTDTFESNPNRSGYTGASCVTACP 300
; 301 YNYLSTDVGSCTLVCPLNQOEVTAEDGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
; 301 YNYLSTDVGSCTLVCPLNQOEVTAEDGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
; 301 YNYLSTDVGSCTLVCPLNQOEVTAEDGTORCEKSKPCARVCYGLGMEHLREVRVTSAN 360
; 361 IOEFAGCKKIFGSLAFLESPFDPDASNTAPLOPEQLQVFETLEITGYLYISAMPDSL 420
; 361 IOEFAGCKKIFGSLAFLESPFDPDASNTAPLOPEQLQVFETLEITGYLYISAMPDSL 420
; 361 IOEFAGCKKIFGSLAFLESPFDPDASNTAPLOPEQLQVFETLEITGYLYISAMPDSL 420
; 421 DLSVQNTQVIRGRILHNGAVSLTQIGISWLGIRSLRELGSGIALIHNHTLCFVHTV 480
; 421 DLSVQNTQVIRGRILHNGAVSLTQIGISWLGIRSLRELGSGIALIHNHTLCFVHTV 480
; 421 DLSVQNTQVIRGRILHNGAVSLTQIGISWLGIRSLRELGSGIALIHNHTLCFVHTV 480
; 481 PMDQFRPHQALHTANRPECEVGEGLACHQLCARHGCMGPGPTQCVNCSQFLRGEC 540
; 481 PMDQFRPHQALHTANRPECEVGEGLACHQLCARHGCMGPGPTQCVNCSQFLRGEC 540
; 481 PMDQFRPHQALHTANRPECEVGEGLACHQLCARHGCMGPGPTQCVNCSQFLRGEC 540
; 541 VEECEVLOGLPREYNARHCLPCHPECOPONGSVTCFGEADQCVACAHYDPPFCVARC 600
; 541 VEECEVLOGLPREYNARHCLPCHPECOPONGSVTCFGEADQCVACAHYDPPFCVARC 600
; 541 VEECEVLOGLPREYNARHCLPCHPECOPONGSVTCFGEADQCVACAHYDPPFCVARC 600
; 601 PSGVPRDLSYMPWKFPEDEBGAQCPPIINCTHSQVLDLDDKGPABQASPLTISIAYVG 660
; 601 PSGVPRDLSYMPWKFPEDEBGAQCPPIINCTHSQVLDLDDKGPABQASPLTISIAYVG 660
; 601 PSGVPRDLSYMPWKFPEDEBGAQCPPIINCTHSQVLDLDDKGPABQASPLTISIAYVG 660
; 661 ILLVVVLGVVGGILIKRQOKIRKTYMRILQETELVPLTPSGAMPQAOQRILKETEL 720
; 661 ILLVVVLGVVGGILIKRQOKIRKTYMRILQETELVPLTPSGAMPQAOQRILKETEL 720
; 661 ILLVVVLGVVGGILIKRQOKIRKTYMRILQETELVPLTPSGAMPQAOQRILKETEL 720

```

```

; 721 RKVVLGSGAFGVYKGIWIPDGENVKI PVAIKYLRENTSPEKANEIIDEAYYVAGVSP 780
; 721 RKVVLGSGAFGVYKGIWIPDGENVKI PVAIKYLRENTSPEKANEIIDEAYYVAGVSP 780
; 721 RKVVLGSGAFGVYKGIWIPDGENVKI PVAIKYLRENTSPEKANEIIDEAYYVAGVSP 780
; 781 YVSRLLGICLSTVQVLTQMLPYGCLLDHYRENNGRIGSODLMMCMQIAGMSVLEBVR 840
; 781 YVSRLLGICLSTVQVLTQMLPYGCLLDHYRENNGRIGSODLMMCMQIAGMSVLEBVR 840
; 781 YVSRLLGICLSTVQVLTQMLPYGCLLDHYRENNGRIGSODLMMCMQIAGMSVLEBVR 840
; 841 LVHRDLAARNVLVSPNHNKITTDFGLARLDIDETEVHADGKVPKIMMALESILRRFT 900
; 841 LVHRDLAARNVLVSPNHNKITTDFGLARLDIDETEVHADGKVPKIMMALESILRRFT 900
; 841 LVHRDLAARNVLVSPNHNKITTDFGLARLDIDETEVHADGKVPKIMMALESILRRFT 900
; 901 HQSDVMSYGVVWMLMFGAPDYDGI PARBLPDLLEKGERLPPPICITDYYMTMVKCM 960
; 901 HQSDVMSYGVVWMLMFGAPDYDGI PARBLPDLLEKGERLPPPICITDYYMTMVKCM 960
; 901 HQSDVMSYGVVWMLMFGAPDYDGI PARBLPDLLEKGERLPPPICITDYYMTMVKCM 960
; 961 IDSECRPRFRELVEFARMARDPQRFVITQNEDEDGPASPLDSTFYRSLLEDDMDGLVDA 1020
; 961 IDSECRPRFRELVEFARMARDPQRFVITQNEDEDGPASPLDSTFYRSLLEDDMDGLVDA 1020
; 961 IDSECRPRFRELVEFARMARDPQRFVITQNEDEDGPASPLDSTFYRSLLEDDMDGLVDA 1020
; 1021 EBYLVPOQGFPCPDPAAGAGMVHRRSSSTRSGGDLTLGLPSEBEARSPPLAPSEG 1080
; 1021 EBYLVPOQGFPCPDPAAGAGMVHRRSSSTRSGGDLTLGLPSEBEARSPPLAPSEG 1080
; 1021 EBYLVPOQGFPCPDPAAGAGMVHRRSSSTRSGGDLTLGLPSEBEARSPPLAPSEG 1080
; 1081 AGSDVFDGLGMAKGIQSLPTHDPSPLOKYSRSDPTVPLPSEFDGYVAPLTCSPQERY 1140
; 1081 AGSDVFDGLGMAKGIQSLPTHDPSPLOKYSRSDPTVPLPSEFDGYVAPLTCSPQERY 1140
; 1081 AGSDVFDGLGMAKGIQSLPTHDPSPLOKYSRSDPTVPLPSEFDGYVAPLTCSPQERY 1140
; 1141 NOPVRPOPSPRRGGLPARAPAGATLERPTLSPGKNGVVDPAFGCAVENPEYLTPQ 1200
; 1141 NOPVRPOPSPRRGGLPARAPAGATLERPTLSPGKNGVVDPAFGCAVENPEYLTPQ 1200
; 1141 NOPVRPOPSPRRGGLPARAPAGATLERPTLSPGKNGVVDPAFGCAVENPEYLTPQ 1200
; 1201 GAAPOPHPPPAFDNLVYMDQDPPERGAPSTFGTPTAENPEYLGLDVEY 1255
; 1201 GAAPOPHPPPAFDNLVYMDQDPPERGAPSTFGTPTAENPEYLGLDVEY 1255
; 1201 GAAPOPHPPPAFDNLVYMDQDPPERGAPSTFGTPTAENPEYLGLDVEY 1255

```

RESULT 9
US-08-466-680B-68
Sequence 68, Application US/08466680B

Patent No. 6075122

GENERAL INFORMATION:
Applicant: Cheever, Martin A.

Title of Invention: IMMUNE REACTIVITY TO HER-2/neu PROTEIN

Title of Invention: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:
ADDRESS: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,680B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.448C4

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 68:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-466-680B-68

Query Match 99.9%; Score 6806; DB 2; Length 1255;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 METALCRWGLLALLPFGAASVQVCTGDMKRLPASPETHLDMRLHYOCQVQGNL 60
 DB 1 METALCRWGLLALLPFGAASVQVCTGDMKRLPASPETHLDMRLHYOCQVQGNL 60
 QY 61 ELYLPTNASLFLQDIOEVQGYVLIANQVQVPLQRLIRVGTQLFEDNYALAVLNG 120
 DB 61 ELYLPTNASLFLQDIOEVQGYVLIANQVQVPLQRLIRVGTQLFEDNYALAVLNG 120
 QY 121 DELNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILMKDIFHKNNOLA 180
 DB 121 DELNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILMKDIFHKNNOLA 180
 QY 181 LTLIDTNRBRACHPCSPMKSGRCWGESSEDCOSLTRTYCAGGCARCKGPLPTDCCHQC 240
 DB 181 LTLIDTNRBRACHPCSPMKSGRCWGESSEDCOSLTRTYCAGGCARCKGPLPTDCCHQC 240
 QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPDEGRYTFGASCVTACP 300
 DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPDEGRYTFGASCVTACP 300
 QY 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
 DB 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
 QY 361 IOEPFAGCKKIFGSLAFLESFPGDPASNTAPLOPQLOVFEETLEETGYLISAMPDSL 420
 DB 361 IOEPFAGCKKIFGSLAFLESFPGDPASNTAPLOPQLOVFEETLEETGYLISAMPDSL 420
 QY 421 DLSVFONLOVIRGRILHNGAYSILTOGLGISWLGRLSRLBELSGALIHNNTHLCFVHTV 480
 DB 421 DLSVFONLOVIRGRILHNGAYSILTOGLGISWLGRLSRLBELSGALIHNNTHLCFVHTV 480
 QY 481 PWDOLFRRNHQALLHTANRBEDECVBGLACHQLCARGHCMWGPPTQVNCSSQPLRGQBC 540
 DB 481 PWDOLFRRNHQALLHTANRBEDECVBGLACHQLCARGHCMWGPPTQVNCSSQPLRGQBC 540
 QY 541 VEECVLQGLPREYVNAHCLPCHPECOQNGSVTCFGEADQVACAHYKDPFVCVAC 600
 DB 541 VEECVLQGLPREYVNAHCLPCHPECOQNGSVTCFGEADQVACAHYKDPFVCVAC 600
 QY 601 PEGVNPDLSTYMPIMKRPDEEGACOPCPINCTHSCVDLDDKGPABEGRASPLTISIISAVVG 660
 DB 601 PEGVNPDLSTYMPIMKRPDEEGACOPCPINCTHSCVDLDDKGPABEGRASPLTISIISAVVG 660
 QY 661 ILLVVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
 DB 661 ILLVVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPNOQMRILKETEL 720
 QY 721 RKRYVLGSGAFGTVYKGIWIPDGENYKIPVAIKVLENTSPRANKEIIDEAYVMGVGSP 780
 DB 721 RKRYVLGSGAFGTVYKGIWIPDGENYKIPVAIKVLENTSPRANKEIIDEAYVMGVGSP 780
 QY 781 YVSRLLGLCTSTVQLVQLMYGCILDHVRENRGLSGODLLNMCQIAKMSYLDEVR 840
 DB 781 YVSRLLGLCTSTVQLVQLMYGCILDHVRENRGLSGODLLNMCQIAKMSYLDEVR 840
 QY 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLIDETETHADGKVPILKMWALSLIRRRPT 900
 DB 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLIDETETHADGKVPILKMWALSLIRRRPT 900
 QY 901 HOSDWSGVYTWELMTFGAKPYDGI PAEIPDLLEKGRLLPOPPICITIDVYMIWKKCM 960
 DB 901 HOSDWSGVYTWELMTFGAKPYDGI PAEIPDLLEKGRLLPOPPICITIDVYMIWKKCM 960

QY 961 IDSECRPRERELVSFBSRWARDPORFVYIQNEDLGPASPLDSTFYRSLLEDDMDGLVDA 1020
 DB 961 IDSECRPRERELVSFBSRWARDPORFVYIQNEDLGPASPLDSTFYRSLLEDDMDGLVDA 1020
 QY 1021 EBYLYPQCGFFCPDPAPGAGVHHRSSSTRSGGDLTLCLEPSEEAAPRSPPLAPSG 1080
 DB 1021 EBYLYPQCGFFCPDPAPGAGVHHRSSSTRSGGDLTLCLEPSEEAAPRSPPLAPSG 1080
 QY 1081 AGSDVPDGLGMAKGLQSLPTHPSPLOQRSEDPVPLPSETGYVAPLFCSPQPEYV 1140
 DB 1081 AGSDVPDGLGMAKGLQSLPTHPSPLOQRSEDPVPLPSETGYVAPLFCSPQPEYV 1140
 QY 1141 NOPDVRPOPSPREBGLPAPAPAGATLERPKTLSPGKGVVDVPAFGAVENPEYLTPQ 1200
 DB 1141 NOPDVRPOPSPREBGLPAPAPAGATLERPKTLSPGKGVVDVPAFGAVENPEYLTPQ 1200
 QY 1201 GGAAPQHPHPAPFSPAFDNLXYWDDPPERKAPBSTFKGTPTAENPEYLGDLVPV 1255
 DB 1201 GGAAPQHPHPAPFSPAFDNLXYWDDPPERKAPBSTFKGTPTAENPEYLGDLVPV 1255

RESULT 10

US-09-527-487-2
 ; Sequence 2, Application US/09527487
 ; Patent No. 6528060
 ; APPLICANT: Nicolette, Charles
 ; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
 ; FILE REFERENCE: 126881309200
 ; CURRENT APPLICATION NUMBER: US/09/527,487
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 1255
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-527-487-2

Query Match 99.9%; Score 6806; DB 2; Length 1255;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 METALCRWGLLALLPFGAASVQVCTGDMKRLPASPETHLDMRLHYOCQVQGNL 60
 DB 1 METALCRWGLLALLPFGAASVQVCTGDMKRLPASPETHLDMRLHYOCQVQGNL 60
 QY 61 ELYLPTNASLFLQDIOEVQGYVLIANQVQVPLQRLIRVGTQLFEDNYALAVLNG 120
 DB 61 ELYLPTNASLFLQDIOEVQGYVLIANQVQVPLQRLIRVGTQLFEDNYALAVLNG 120
 QY 121 DELNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILMKDIFHKNNOLA 180
 DB 121 DELNNTTPVTGASPGGLRELQRLSLTEILKGGVLIQRNPOLCYQDTILMKDIFHKNNOLA 180
 QY 181 LTLIDTNRBRACHPCSPMKSGRCWGESSEDCOSLTRTYCAGGCARCKGPLPTDCCHQC 240
 DB 181 LTLIDTNRBRACHPCSPMKSGRCWGESSEDCOSLTRTYCAGGCARCKGPLPTDCCHQC 240
 QY 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPDEGRYTFGASCVTACP 300
 DB 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMNPDEGRYTFGASCVTACP 300
 QY 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
 DB 301 YNYLSTDVGSCTLVCPPLHNOEVTABDGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
 QY 361 IOEPFAGCKKIFGSLAFLESFPGDPASNTAPLOPQLOVFEETLEETGYLISAMPDSL 420
 DB 361 IOEPFAGCKKIFGSLAFLESFPGDPASNTAPLOPQLOVFEETLEETGYLISAMPDSL 420
 QY 421 DLSVFONLOVIRGRILHNGAYSILTOGLGISWLGRLSRLBELSGALIHNNTHLCFVHTV 480

```
Db 421 DLSVFQNIQVIRGRILHNGAVSLTLOGLISWLGSRSLRELSGSLALIHNTHTLCFVHTV 480
Qy 481 PMDQLFRRPHQALHTANRPEDECVGEGLACHQLCARGHCMGPGPTQCVNCSQFLRGEC 540
Db 481 PMDQLFRRPHQALHTANRPEDECVGEGLACHQLCARGHCMGPGPTQCVNCSQFLRGEC 540
Qy 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCFGEADQCYACAHYKDPPFCVARC 600
Db 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCFGEADQCYACAHYKDPPFCVARC 600
Qy 601 PSGVPRDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDDKCCPAEQRASPLTSISAVVG 660
Db 601 PSGVPRDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDDKCCPAEQRASPLTSISAVVG 660
Qy 661 ILVVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPQAOQRILIKETEL 720
Db 661 ILVVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPQAOQRILIKETEL 720
Qy 721 RKVVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPYANKELIDEAVYMAVGSP 780
Db 721 RKVVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPYANKELIDEAVYMAVGSP 780
Qy 781 YVSRLLGICLTSTVQLVQLMFYGCLLDHVRENRGLSGODLWMCQIAGKMSYLEBVR 840
Db 781 YVSRLLGICLTSTVQLVQLMFYGCLLDHVRENRGLSGODLWMCQIAGKMSYLEBVR 840
Qy 841 LVHRDLAARNVLVKSPPNHVKITDPGLARLLIDETEVHADGKVPDKMMALESIIRRRFT 900
Db 841 LVHRDLAARNVLVKSPPNHVKITDPGLARLLIDETEVHADGKVPDKMMALESIIRRRFT 900
Qy 901 HQSDVMSGYTVWELMTFGAKPYDGI PARFIPDLLEKGRLLPQPICTIDVYMIWVKCM 960
Db 901 HQSDVMSGYTVWELMTFGAKPYDGI PARFIPDLLEKGRLLPQPICTIDVYMIWVKCM 960
Qy 961 IDSECRPRPRELVSEFSRMARDPQRFVVIQNEDELGPASPLDSTFYRSILLEDDMGDLVDA 1020
Db 961 IDSECRPRPRELVSEFSRMARDPQRFVVIQNEDELGPASPLDSTFYRSILLEDDMGDLVDA 1020
Qy 1021 EBYLVPOQGFPCPDPAAPAGVYHRRHSSSTRSGGDLTLGLBESSEBAPASPLAPSEG 1080
Db 1021 EBYLVPOQGFPCPDPAAPAGVYHRRHSSSTRSGGDLTLGLBESSEBAPASPLAPSEG 1080
Qy 1081 AGSDVFDGDLGMAKGLQSLPTHDPSLQRYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140
Db 1081 AGSDVFDGDLGMAKGLQSLPTHDPSLQRYSEDPVPLPSETDGYVAPLTCSPQPEYV 1140
Qy 1141 NOPDVRPQPSBREGPLPAAPAGATLERPKTSLSGKGVYKDVAPFGAVERNPEYLTPO 1200
Db 1141 NOPDVRPQPSBREGPLPAAPAGATLERPKTSLSGKGVYKDVAPFGAVERNPEYLTPO 1200
Qy 1201 GGAAPQPHPPAPFAPFNLVYWDQPPERGAPESTFGKTPTAENPEYLGLDVPI 1255
Db 1201 GGAAPQPHPPAPFAPFNLVYWDQPPERGAPESTFGKTPTAENPEYLGLDVPI 1255

RESULT 11
US-09-811-115-3
; Sequence 3, Application us/09811115
; Patent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENEENT 034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1255
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3
Query Match 99.8%; Score 6806; DB 2; Length 1255;
Best Local Similarity 99.8%; Fred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MELAALCMWGLLMLALPPGAASVQCTGTDKMLRLPASPETHLDMLRHLYOGCCQVVGNTL 60
Db 1 MELAALCMWGLLMLALPPGAASVQCTGTDKMLRLPASPETHLDMLRHLYOGCCQVVGNTL 60
Qy 61 ELYTLPINASISPLDIOEVQGYVLIANQVRQVPLQRLIRYKGTQLFEDNYALAVLNG 120
Db 61 ELYTLPINASISPLDIOEVQGYVLIANQVRQVPLQRLIRYKGTQLFEDNYALAVLNG 120
Qy 121 DPLNNTTPTVTAASPGGLRELOLRSLTEILKGGVLIQRNPOLCYQDTILMKDIFKNNOLA 180
Db 121 DPLNNTTPTVTAASPGGLRELOLRSLTEILKGGVLIQRNPOLCYQDTILMKDIFKNNOLA 180
Qy 181 LTLIDTNRSRACHPCSFWCKSGSRGESSBDCQSLTRTVCAAGCARCKGPLPTDCHBQC 240
Db 181 LTLIDTNRSRACHPCSFWCKSGSRGESSBDCQSLTRTVCAAGCARCKGPLPTDCHBQC 240
Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNDEGRTTGASCVTACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNDEGRTTGASCVTACP 300
Qy 301 YNYLSTDVSGCTIACPLHNOEVTAEQDQCEKSKPCARCYGIGMHLREVRVTSAN 360
Db 301 YNYLSTDVSGCTIACPLHNOEVTAEQDQCEKSKPCARCYGIGMHLREVRVTSAN 360
Qy 361 IOEFAGCKKIFGSLAFIPESFDGPASNTABLOEQLOVFETLEITGYLIISAMPDSL 420
Db 361 IOEFAGCKKIFGSLAFIPESFDGPASNTABLOEQLOVFETLEITGYLIISAMPDSL 420
Qy 421 DLSVFQNIQVIRGRILHNGAVSLTLOGLISWLGSRSLRELSGSLALIHNTHTLCFVHTV 480
Db 421 DLSVFQNIQVIRGRILHNGAVSLTLOGLISWLGSRSLRELSGSLALIHNTHTLCFVHTV 480
Qy 481 PMDQLFRRPHQALHTANRPEDECVGEGLACHQLCARGHCMGPGPTQCVNCSQFLRGEC 540
Db 481 PMDQLFRRPHQALHTANRPEDECVGEGLACHQLCARGHCMGPGPTQCVNCSQFLRGEC 540
Qy 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCFGEADQCYACAHYKDPPFCVARC 600
Db 541 VEECRVLOGLPREVYNARHCLPCHPECOPONGSVTCFGEADQCYACAHYKDPPFCVARC 600
Qy 601 PSGVPRDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDDKCCPAEQRASPLTSISAVVG 660
Db 601 PSGVPRDLSYMPIMKFPDEEGACQPCPINCTHSCVDLDDKCCPAEQRASPLTSISAVVG 660
Qy 661 ILVVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPQAOQRILIKETEL 720
Db 661 ILVVVLGVVFGILLIKRQOKIRKTYMRLLQETELVEPLTPSGAMPQAOQRILIKETEL 720
Qy 721 RKVVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPYANKELIDEAVYMAVGSP 780
Db 721 RKVVLGSGAGFTVYKGIWIPDGENVKIPVAIKVIRENTSPYANKELIDEAVYMAVGSP 780
Qy 781 YVSRLLGICLTSTVQLVQLMFYGCLLDHVRENRGLSGODLWMCQIAGKMSYLEBVR 840
Db 781 YVSRLLGICLTSTVQLVQLMFYGCLLDHVRENRGLSGODLWMCQIAGKMSYLEBVR 840
Qy 841 LVHRDLAARNVLVKSPPNHVKITDPGLARLLIDETEVHADGKVPDKMMALESIIRRRFT 900
Db 841 LVHRDLAARNVLVKSPPNHVKITDPGLARLLIDETEVHADGKVPDKMMALESIIRRRFT 900
Qy 901 HQSDVMSGYTVWELMTFGAKPYDGI PARFIPDLLEKGRLLPQPICTIDVYMIWVKCM 960
Db 901 HQSDVMSGYTVWELMTFGAKPYDGI PARFIPDLLEKGRLLPQPICTIDVYMIWVKCM 960
Qy 961 IDSECRPRPRELVSEFSRMARDPQRFVVIQNEDELGPASPLDSTFYRSILLEDDMGDLVDA 1020
```


Db 961 IDSECRPRFRELVSFSSMARDPQRFVYIQNEDLPASPGLDSTFYSRLLEDMDMDLVDA 1020
 QY 1021 BEYLVPOGFCPPDPAFGAGVWHHRHSSSTRSGGDTLLGLBSEBEPAPSLAPSBG 1080
 Db 1021 BEYLVPOGFCPPDPAFGAGVWHHRHSSSTRSGGDTLLGLBSEBEPAPSLAPSBG 1080
 QY 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYAAPLCSPOPEVY 1140
 Db 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYAAPLCSPOPEVY 1140
 QY 1141 NOPDVRPOPSPREBEPPLPAAPAGATLERPKTLSPGKGVVDVAFGAVENPEYLTPO 1200
 Db 1141 NOPDVRPOPSPREBEPPLPAAPAGATLERPKTLSPGKGVVDVAFGAVENPEYLTPO 1200
 QY 1201 GGAAPQHPPPAFSPAFDNLVYWDOPPERGAPPESTFKGTPTAENPEYLGDLVPV 1255
 Db 1201 GGAAPQHPPPAFSPAFDNLVYWDOPPERGAPPESTFKGTPTAENPEYLGDLVPV 1255

RESULT 12

US-09-354-533-68

Sequence 68, Application US/09354533

Patent No. 6664370

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

Diale, Mary L.

TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN

FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE

HER-2/neu ONCOGENE IS ASSOCIATED

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSER: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/354,533

FILING DATE: 15-Jul-1999

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.448C9

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 68:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 68:

us-09-354-533-68

Query Match

Best Local Similarity 99.9%; Score 6806; DB 2; Length 1255;

Matches 1554; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 METALCRKGLLALPFGAASCTGCTDMKLRLLPASPEETHLMDLRHLVYSCQVVGSL 60
 Db 1 METALCRKGLLALPFGAASCTGCTDMKLRLLPASPEETHLMDLRHLVYSCQVVGSL 60
 QY 61 BLTYLPTNASLSFLDIOEVGVYLIANQVROVPLQRLIVRGTLQPEEDNYALAVLNG 120
 Db 61 BLTYLPTNASLSFLDIOEVGVYLIANQVROVPLQRLIVRGTLQPEEDNYALAVLNG 120

QY 121 DPLANTTAVTGASPGGLRELOLRSLTBILKGVLIQRNPOLCYOPTILMKOIFHNQOLA 180
 Db 121 DPLANTTAVTGASPGGLRELOLRSLTBILKGVLIQRNPOLCYOPTILMKOIFHNQOLA 180
 QY 181 LTLIDTNRSRACHPCSPCKSGRCMGSSSDCQSLTRTVACGACRCAGPLPTDCCHQC 240
 Db 181 LTLIDTNRSRACHPCSPCKSGRCMGSSSDCQSLTRTVACGACRCAGPLPTDCCHQC 240
 QY 241 AAGCTGPGHSDCLACLHFNHSGICELHCPALVTYNTDFESMPNBEGRVTFGASCVTACP 300
 Db 241 AAGCTGPGHSDCLACLHFNHSGICELHCPALVTYNTDFESMPNBEGRVTFGASCVTACP 300
 QY 301 YNLTSDVSGCTLVCPLEHNOEYTAEDGTORCEKSPCARVCYGLGMEHLREVRVTSAN 360
 Db 301 YNLTSDVSGCTLVCPLEHNOEYTAEDGTORCEKSPCARVCYGLGMEHLREVRVTSAN 360
 QY 361 IOEFAGCKKIFGSLAFPLESPFDGDPASNTAPLQPELOVFTLEETGYLIISAMPDILP 420
 Db 361 IOEFAGCKKIFGSLAFPLESPFDGDPASNTAPLQPELOVFTLEETGYLIISAMPDILP 420
 QY 421 DLSVFQNTQVTRGRILHNGAYSLLTQIGISWLGSRSLRELSSGLALIHNTHLCEVHTV 480
 Db 421 DLSVFQNTQVTRGRILHNGAYSLLTQIGISWLGSRSLRELSSGLALIHNTHLCEVHTV 480
 QY 481 PMDOLFRPHOALLHTANRPEDECVGEGIAHQLCARGHCMGPPTQCVCNCSQFLRGQC 540
 Db 481 PMDOLFRPHOALLHTANRPEDECVGEGIAHQLCARGHCMGPPTQCVCNCSQFLRGQC 540
 QY 541 VEBECVLOGLPREVYNAHCLPCHBECOPQNGSVTCFGEADQCVACAHYKPPFCVARC 600
 Db 541 VEBECVLOGLPREVYNAHCLPCHBECOPQNGSVTCFGEADQCVACAHYKPPFCVARC 600
 QY 601 PSQVPRDLSYMPYKFPPEBEGACOPCPINCHTSCVDLDDKGRAPORASPLTIIISAVG 660
 Db 601 PSQVPRDLSYMPYKFPPEBEGACOPCPINCHTSCVDLDDKGRAPORASPLTIIISAVG 660
 QY 661 ILVVVLGVVFGILIKRQOKIRKTYMRLLQETELVEPLFPGAMPNOQMRILKETEL 720
 Db 661 ILVVVLGVVFGILIKRQOKIRKTYMRLLQETELVEPLFPGAMPNOQMRILKETEL 720
 QY 721 RKVKVLSGARGTYVKGIMIPDGENVKI PVAKVLRNTSPRANKELIDEAYVMAVGSP 780
 Db 721 RKVKVLSGARGTYVKGIMIPDGENVKI PVAKVLRNTSPRANKELIDEAYVMAVGSP 780
 QY 781 YVSRLLGICLTSTVQVLTQMLPYGCLLHVNRGRSLQSDLLMCMQIAKMSYLEBYR 840
 Db 781 YVSRLLGICLTSTVQVLTQMLPYGCLLHVNRGRSLQSDLLMCMQIAKMSYLEBYR 840
 QY 841 LVHRDLAARNVLYVSPNHYKITDPGLARLDIDETEHADGKVPDKMMALESIIRRP 900
 Db 841 LVHRDLAARNVLYVSPNHYKITDPGLARLDIDETEHADGKVPDKMMALESIIRRP 900
 QY 901 HQSDVMSYGVTVWELMTFGAKPYDGI PARHLPDLLEKBERLPQPICTIDVYMIWVKCM 960
 Db 901 HQSDVMSYGVTVWELMTFGAKPYDGI PARHLPDLLEKBERLPQPICTIDVYMIWVKCM 960
 QY 961 IDSECRPPRRELVSFSSMARDPQRFVYIQNEDLPASPGLDSTFYSRLLEDMDMDLVDA 1020
 Db 961 IDSECRPPRRELVSFSSMARDPQRFVYIQNEDLPASPGLDSTFYSRLLEDMDMDLVDA 1020
 QY 1021 BEYLVPOGFCPPDPAFGAGVWHHRHSSSTRSGGDTLLGLBSEBEPAPSLAPSBG 1080
 Db 1021 BEYLVPOGFCPPDPAFGAGVWHHRHSSSTRSGGDTLLGLBSEBEPAPSLAPSBG 1080
 QY 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYAAPLCSPOPEVY 1140
 Db 1081 AGSDVFDGDLGMAKAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYAAPLCSPOPEVY 1140
 QY 1141 NOPDVRPOPSPREBEPPLPAAPAGATLERPKTLSPGKGVVDVAFGAVENPEYLTPO 1200
 Db 1141 NOPDVRPOPSPREBEPPLPAAPAGATLERPKTLSPGKGVVDVAFGAVENPEYLTPO 1200

Qy 1201 GGAAPQHPHPAPSPAFNDLYWDDPPERGAPSTFKGPTPAENPEYLGLDVPY 1255
Db 1201 GGAAPQHPHPAPSPAFNDLYWDDPPERGAPSTFKGPTPAENPEYLGLDVPY 1255

RESULT 13

US-08-484-438-8
Sequence 8, Application US/08484438
Patent No. 5811098
Patent No. 5811098 5780031
GENERAL INFORMATION:
APPLICANT: Plozman, Gregory D.
APPLICANT: Culouscou, Jean-Michel
APPLICANT: Shoyab, Mohammed
APPLICANT: Siegal, Clay B.
APPLICANT: Helictr m, Ingegerd
APPLICANT: Helictr m, Karl E.
TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-8

Query Match 98.8%; Score 6732; DB 1; Length 1255;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1247; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

Qy 1 MELALCWMGLLALLPFGAASCTGCTDMKRLRPASPTHLDMLRHLYOGCCQVVGNTL 60
Db 1 MELALCWMGLLALLPFGAASCTGCTDMKRLRPASPTHLDMLRHLYOGCCQVVGNTL 60

Qy 61 ELTYLPTNASSLFLODIOEVQGYVLIANQVQVPLQRLRYRGQLTFEDNYAALVLDNG 120
Db 61 ELTYLPTNASSLFLODIOEVQGYVLIANQVQVPLQRLRYRGQLTFEDNYAALVLDNG 120

Qy 121 DPLNNTTPVTGASPGGLRELQLRSLEILKGGVLIQSNPOLCYODTILMKDI FHKNNOLA 180
Db 121 DPLNNTTPVTGASPGGLRELQLRSLEILKGGVLIQSNPOLCYODTILMKDI FHKNNOLA 180

Qy 181 LTLIDTNRSRACHCSPCKGSRGWSGSSSDCCSLTTRVCAGGACRCKGRLPTCCCHQC 240
Db 181 LTLIDTNRSRACHCSPCKGSRGWSGSSSDCCSLTTRVCAGGACRCKGRLPTCCCHQC 240

Qy 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCTYACP 300
Db 241 AAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPEGRTYTGASCTYACP 300

Qy 301 YNYLSTDVGSCCTVCPHNOEVTAE DSTGRCCKSKPCARVCYGLGMEHLREVAVTSAN 360
Db 301 YNYLSTDVGSCCTVCPHNOEVTAE DSTGRCCKSKPCARVCYGLGMEHLREVAVTSAN 360

Qy 361 IOEPAGCKKIFGSLAPFESFDGDPASNTAPLOEBOQVFPETLEITGYLISAMPISLP 420
Db 361 IOEPAGCKKIFGSLAPFESFDGDPASNTAPLOEBOQVFPETLEITGYLISAMPISLP 420

Qy 421 DLSVFQNLQVYRGRILHNGAYSLTLOGLSWLGRLSRLRELSGLALIHNTTHCFVYTV 480
Db 421 DLSVFQNLQVYRGRILHNGAYSLTLOGLSWLGRLSRLRELSGLALIHNTTHCFVYTV 480

Qy 481 PMDQLFNPHQALHTANRPEDECVGEGLAHQLCARGHCWGPPTQCVNCSQFLRQEC 540
Db 481 PMDQLFNPHQALHTANRPEDECVGEGLAHQLCARRALLGSGPTQCVNCSQFLRQEC 540

Qy 541 VEECRVYQGLPREVYVNAHCLCPHCEQOPONGSTTCGPEADOCVACHYNDPPFCVARC 600
Db 541 VEECRVYQGLPREVYVNAHCLCPHCEQOPONGSTTCGPEADOCVACHYNDPPFCVARC 600

Qy 601 PSQVCPDLSYMPYKPFDEBACQPCPINCCHSCVDLDDKCPMBQASPLTSLVSNVVG 660
Db 601 PSQVCPDLSYMPYKPFDEBACQPCPINCCHSCVDLDDKCPMBQASPLTSLVSNVVG 660

Qy 720 ILVVLGVVFGILIKRQOKIRRYTMRRLIQETELVEPLTPSGAMPNQAOMRLKETEL 720
Db 720 ILVVLGVVFGILIKRQOKIRRYTMRRLIQETELVEPLTPSGAMPNQAOMRLKETEL 720

Qy 780 RKVAVLGSSAGCTYTKGIWIPDGENVKIPVAIKYLRNTSPKAKETILDEAYVNAAGVSP 780
Db 721 RKVAVLGSSAGCTYTKGIWIPDGENVKIPVAIKYLRNTSPKAKETILDEAYVNAAGVSP 780

Qy 840 YVSRLLGICLSTYQVLTQMLPYGCLLDHYENRGRLGSDQLLNMCMQIAKMSYLEDVR 840
Db 781 YVSRLLGICLSTYQVLTQMLPYGCLLDHYENRGRLGSDQLLNMCMQIAKMSYLEDVR 840

Qy 900 LVHBDLAARNVLYKSPNHVKITDFGLARLLDIDETRYHADGKVPYKMALESILRRFT 900
Db 841 LVHBDLAARNVLYKSPNHVKITDFGLARLLDIDETRYHADGKVPYKMALESILRRFT 900

Qy 960 HQSDVMSGYTVWMLMTPGAKPYDGIPIAREIPDLLEGERLPORPCTIDVYMMVMVCKM 960
Db 901 HQSDVMSGYTVWMLMTPGAKPYDGIPIAREIPDLLEGERLPORPCTIDVYMMVMVCKM 960

Qy 1020 IDSECRPRFRELVESEFARMADPQRFVYIQNEBGLPASPLDSTFYRSLLEDDMGDLVDA 1020
Db 961 IDSECRPRFRELVESEFARMADPQRFVYIQNEBGLPASPLDSTFYRSLLEDDMGDLVDA 1020

Qy 1080 BEYLVPQGFPCDPAPGAGGVHRRSSSTRSGGDLTYGLEPSEBAPRSPLAPSEG 1080
Db 1021 BEYLVPQGFPCDPAPGAGGVHRRSSSTRSGGDLTYGLEPSEBAPRSPLAPSEG 1080

Qy 1140 AGSVFPGDGLMGAKAGIQSLPTHDPSPLOQYSSDDPVYPLSEFDGYVAPLTCSPPQPEYV 1140
Db 1081 AGSVFPGDGLMGAKAGIQSLPTHDPSPLOQYSSDDPVYPLSEFDGYVAPLTCSPPQPEYV 1140

Qy 1200 NQPDVRPOPSPESRGSLPAAPAGATLERPTLSPGKGVVVDYAFAGAVENPEYLTPO 1200
Db 1141 NQPDVRPOPSPESRGSLPAAPAGATLERPTLSPGKGVVVDYAFAGAVENPEYLTPO 1200

[illegible]

RESULT 14
US-09-146-283-4
Sequence 4, Application US/09146283

GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Ruegg, Curtis L.
APPLICANT: Wu, Hongyu
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/146,283
FILING DATE: 03-SEPT-1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:

Query Match	53.3%	Score 3630	DB 1	Length 782
Best Local Similarity	99.8%	Pred. No. 5.9e-238		
Matches 653	Conservative	1	Mismatches 0	Indels 0

QY	MEIATAACRMGLLALALLPFGAASQVCTGDMTLRLPAPERTHLDMLRLHYGGCQVAGNL	60
QY <td>1</td> <td>60</td>	1	60
Db	1	60
QY	61	120
Db	61	120
QY	121	180
Db	121	180
QY	181	240
Db	181	240
QY	241	300
Db	241	300

Qy	301	YNLSTDVSCITLVCJLHNOEYTAEDGFORCEKSKPCARCYGJGMEHLREVAATVSAN	360
Db	301	YNLSTDVSCITLVCJLHNOEYTAEDGFORCEKSKPCARCYGJGMEHLREVAATVSAN	360
Qy	361	IOEFACCKKIFGSLAFLPESFPGDASWTAPLOPEOLQVFETLEETIGYLYISAMPDPL	420
Db	361	IOEFACCKKIFGSLAFLPESFPGDASWTAPLOPEOLQVFETLEETIGYLYISAMPDPL	420
Qy	421	DLISVFNOLQVIRGRILHNGAVSLTLQGLGISWLGISRLREJAGSLALIHNTHLCPVHTV	480
Db	421	DLISVFNOLQVIRGRILHNGAVSLTLQGLGISWLGISRLREJAGSLALIHNTHLCPVHTV	480
Qy	481	PMQULFRNPHQALHTNANRPEBCEYGBEGJACHQLCARHCWGPBPTQVNCVSQFLRQEC	540
Db	481	PMQULFRNPHQALHTNANRPEBCEYGBEGJACHQLCARHCWGPBPTQVNCVSQFLRQEC	540
Qy	541	VEBCRYTLQGLPREYVNAHCLPCHBECOPONGSVYCFEPBADDQCVACAHYKDPFCVARC	600
Db	541	VEBCRYTLQGLPREYVNAHCLPCHBECOPONGSVYCFEPBADDQCVACAHYKDPFCVARC	600
Qy	601	PSGVKCDLSYMPWKRPDEGACQCPINCTHSCVLDLDDKCPAEBORASPLTSL	654
Db	601	PSGVKCDLSYMPWKRPDEGACQCPINCTHSCVLDLDDKCPAEBORASPLTSL	654

RESULT 15
US-08-579-823A-4
; Sequence 4, Application US/08579823A

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

?      COMPUTER READABLE FORM:
?      MEDIUM TYPE: Floppy disk
?      COMPUTER: IBM PC compatible
?      OPERATING SYSTEM: PC-DOS/MS-DOS
?      SOFTWARE: Patent In Release #1.0, Version #1.25
?      CURRENT APPLICATION DATA:
?      CURRENT AMBOND: HC/06/570 003

```

Query Match	Score	DB 2;	Length
Best Local Similarity	Pred. No.	5.9e-238;	
53.3%	3630;		782;
99.8%;			

```
Matches 653; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELAALCRMGLIALLPFGAASSTOVCTGDMKLRLPASPEETHLDMRLHYOGCQVVGNTL 60
Db 1 MELAALCRMGLIALLPFGAASSTOVCTGDMKLRLPASPEETHLDMRLHYOGCQVVGNTL 60
QY 61 ELTYLPTNASLSFLQDIOEVQGVYLIANHVRQVPLQRLRIYRGTOLEFDNYALAVLDNG 120
Db 61 ELTYLPTNASLSFLQDIOEVQGVYLIANHVRQVPLQRLRIYRGTOLEFDNYALAVLDNG 120
QY 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTIIMKDIFFKNQOLA 180
Db 121 DPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPQLCYQDTIIMKDIFFKNQOLA 180
QY 181 LTLIDPTNSRACHPCSPMCKSGRCWGESSEDCOSLTRTVACAGCARCKGPLPTDCHEQC 240
Db 181 LTLIDPTNSRACHPCSPMCKSGRCWGESSEDCOSLTRTVACAGCARCKGPLPTDCHEQC 240
QY 241 AAGCTGPRHSDCLACLFHNHSGICEIACPALVTYNTDTFESMPNDEGRYTFGASCTIACP 300
Db 241 AAGCTGPRHSDCLACLFHNHSGICEIACPALVTYNTDTFESMPNDEGRYTFGASCTIACP 300
QY 301 YNYLSTDVSGCTLVCEPLHNOEVTAEQTCRCEKSKPCARVCYGLGMEHLREVRATVSAN 360
Db 301 YNYLSTDVSGCTLVCEPLHNOEVTAEQTCRCEKSKPCARVCYGLGMEHLREVRATVSAN 360
QY 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVFETLEBITGYLYISAMPDSL 420
Db 361 IOEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVFETLEBITGYLYISAMPDSL 420
QY 421 DLSVFQNTQVIRGRILHNGAVSLTQGLGISWLGLSLRELGSGIALIHNTHTLCFVHTV 480
Db 421 DLSVFQNTQVIRGRILHNGAVSLTQGLGISWLGLSLRELGSGIALIHNTHTLCFVHTV 480
QY 481 PMDQLFRRNPQALLHTANRPEDECVGEGIACHQLCARGHCWGPPTCCVNCOSQFLRGQEC 540
Db 481 PMDQLFRRNPQALLHTANRPEDECVGEGIACHQLCARGHCWGPPTCCVNCOSQFLRGQEC 540
QY 541 VEECEVLQGLPREYVNAHCHLPCHEQCQPNQNGSVTCFGEADQCVACAHYKDPFCVARC 600
Db 541 VEECEVLQGLPREYVNAHCHLPCHEQCQPNQNGSVTCFGEADQCVACAHYKDPFCVARC 600
QY 601 PSGVKPDLISYMPIWKFPDEEGACQPCPINCTHSCVDLDDKGCAPAFQRASPLTSTI 654
Db 601 PSGVKPDLISYMPIWKFPDEEGACQPCPINCTHSCVDLDDKGCAPAFQRASPLTSTL 654
```

Search completed: January 3, 2006, 11:18:33
Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2006, 11:16:48 ; Search time 13 Seconds
(without alignments)
722.957 Million cell updates/sec

Title: US-09-930-125-2
Perfect score: 6815
Sequence: 1 MELALCRWGLLALPPGA.....TFKGTPTAENPEYLQDVPV 1255

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 57103 seqs, 7488799 residues

Total number of hits satisfying chosen parameters: 57103

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA New:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
4: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
5: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB pep.*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB pep.*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*

Printed. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6815	100.0	1255	US-11-022-562-213	Sequence 213, App
2	6806	99.9	1255	US-10-770-726-62	Sequence 62, App1
3	6806	99.9	1255	US-11-113-202-10	Sequence 10, App1
4	3167	46.5	1210	US-11-113-202-6	Sequence 6, App1
5	3167	46.5	1210	US-11-145-566-1	Sequence 1, App1
6	3004.5	44.1	1308	US-11-113-202-16	Sequence 16, App1
7	2545.5	37.4	943	US-11-113-202-8	Sequence 8, App1
8	2441.5	35.8	1342	US-10-770-726-63	Sequence 63, App1
9	2441.5	35.8	1342	US-11-113-202-12	Sequence 12, App1
10	2441.5	35.8	1342	US-11-113-202-14	Sequence 14, App1
11	1874.5	27.5	419	US-11-113-202-23	Sequence 23, App1
12	1873.5	27.5	419	US-11-113-202-4	Sequence 4, App1
13	1860.5	27.3	419	US-11-113-202-2	Sequence 2, App1
14	1056.5	15.5	534	US-11-077-386-25	Sequence 25, App1
15	798	11.7	400	US-11-077-386-26	Sequence 26, App1
16	734	10.8	346	US-11-077-386-24	Sequence 24, App1
17	651	9.6	1367	US-10-995-561-538	Sequence 538, App
18	651	9.6	1367	US-11-145-202-18	Sequence 18, App1
19	650.5	9.5	1368	US-10-995-561-539	Sequence 539, App
20	607	8.9	987	US-10-949-720-395	Sequence 395, App
21	606	8.9	990	US-10-821-234-1201	Sequence 1201, App
22	584.5	8.6	976	US-10-966-483-2	Sequence 2, App1
23	583.5	8.6	1035	US-10-966-483-20	Sequence 20, App1
24	583.5	8.6	1035	US-11-021-441-4	Sequence 4, App1
25	582.5	8.5	984	US-11-113-424-60	Sequence 60, App1

26	560	8.2	456	6	US-10-966-483-34	Sequence 34, App1
27	560	8.2	456	7	US-11-021-441-18	Sequence 18, App1
28	552.5	8.1	985	7	US-11-113-424-61	Sequence 61, App1
29	551	8.1	998	6	US-10-510-524-1	Sequence 1, App1
30	543.5	8.0	987	6	US-10-770-726-61	Sequence 61, App1
31	542	8.0	879	6	US-10-770-726-78	Sequence 78, App1
32	542	8.0	995	7	US-11-113-424-62	Sequence 62, App1
33	528	7.7	479	6	US-10-966-483-36	Sequence 36, App1
34	528	7.7	479	7	US-11-021-441-20	Sequence 20, App1
35	528	7.7	490	6	US-10-966-483-42	Sequence 42, App1
36	528	7.7	490	7	US-11-021-441-26	Sequence 26, App1
37	528	7.7	497	6	US-10-966-483-38	Sequence 38, App1
38	528	7.7	497	6	US-10-966-483-40	Sequence 40, App1
39	528	7.7	497	7	US-11-021-441-22	Sequence 22, App1
40	528	7.7	497	7	US-11-021-441-24	Sequence 24, App1
41	526	7.7	983	7	US-11-113-424-59	Sequence 59, App1
42	526	7.6	1005	7	US-11-113-424-63	Sequence 63, App1
43	505	7.4	293	7	US-11-092-168-11	Sequence 11, App1
44	493.5	7.2	1035	7	US-11-113-424-20	Sequence 20, App1
45	492	7.2	1052	7	US-11-113-424-22	Sequence 22, App1

ALIGNMENTS

RESULT 1									
US-11-022-562-213									
Sequence 213, Application US/11022562									
Publication No. US20050249742A1									
GENERAL INFORMATION:									
APPLICANT: Shuang, Ruth M.									
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING									
FILE REFERENCE: DEN-042CN									
CURRENT APPLICATION NUMBER: US/11/022,562									
CURRENT FILING DATE: 2004-12-22									
PCT APPLICATION NUMBER: PCT/US03/20322									
PCT FILING DATE: 2003-06-27									
PRIOR APPLICATION NUMBER: 60/392718									
PRIOR FILING DATE: 2002-06-27									
NUMBER OF SEQ ID NOS: 340									
SOFTWARE: FastSeq for Windows Version 4.0									
SEQ ID NO 213									
LENGTH: 1255									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-11-022-562-213									
Query Match	100.0%	Score 6815	DB 7	Length 1255					
Best Local Similarity	100.0%	Pred. No. 0							
Matches 1255	Conservative 0	Mismatches 0	Indels 0	Gaps 0					
1	MELALCRWGLLALPPGA	STGYCTDKRLPASPETHLDMRLHYGCGVVGNTL	60						
1	MELALCRWGLLALPPGA	STGYCTDKRLPASPETHLDMRLHYGCGVVGNTL	60						
61	ELTILPTASISFLDIO	EVGYVLIANQVROPLOLRIVRGQLFEDNYALAVLNG	120						
61	ELTILPTASISFLDIO	EVGYVLIANQVROPLOLRIVRGQLFEDNYALAVLNG	120						
61	ELTILPTASISFLDIO	EVGYVLIANQVROPLOLRIVRGQLFEDNYALAVLNG	120						
121	DLPLNTTVPVTPAS	FGRLRLTSLTELRKGVLIQRPOLCYOPTIMKDIFFKNNOLA	180						
121	DLPLNTTVPVTPAS	FGRLRLTSLTELRKGVLIQRPOLCYOPTIMKDIFFKNNOLA	180						
121	DLPLNTTVPVTPAS	FGRLRLTSLTELRKGVLIQRPOLCYOPTIMKDIFFKNNOLA	180						
181	LTLIDTNRSPACHP	CPCKSCRCWGSSESDCSLITRVACGACRCKGRLPTDCBQC	240						
181	LTLIDTNRSPACHP	CPCKSCRCWGSSESDCSLITRVACGACRCKGRLPTDCBQC	240						
181	LTLIDTNRSPACHP	CPCKSCRCWGSSESDCSLITRVACGACRCKGRLPTDCBQC	240						
241	AAAGCTPGHSGC	ACLPFHNSGICELHCPALVYNTDTFESMNPBEGRYTGAACVTAACP	300						
241	AAAGCTPGHSGC	ACLPFHNSGICELHCPALVYNTDTFESMNPBEGRYTGAACVTAACP	300						
241	AAAGCTPGHSGC	ACLPFHNSGICELHCPALVYNTDTFESMNPBEGRYTGAACVTAACP	300						

QY 301 YNYLSTDVGSCTIVCPPLHNOEVTADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
DB 301 YNYLSTDVGSCTIVCPPLHNOEVTADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
QY 361 IOEPAGCKKITGSLAFPLPESFDGPASNTAPLOPQOLOVFTLEITGYLYISAMPDLP 420
DB 361 IOEPAGCKKITGSLAFPLPESFDGPASNTAPLOPQOLOVFTLEITGYLYISAMPDLP 420
QY 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISMGLRSRLRELSGSLALIHNTHTLCFYVTV 480
DB 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISMGLRSRLRELSGSLALIHNTHTLCFYVTV 480
QY 481 PMDQLFRNPQALHTANRPEDECVBGLACHOLCARGHCWGPPTQCVCNCSQFLRGQEC 540
DB 481 PMDQLFRNPQALHTANRPEDECVBGLACHOLCARGHCWGPPTQCVCNCSQFLRGQEC 540
QY 541 VEECRVLOGLPREVYNARHCLPCHPECOPOGNSVTCFGEADOCVACHYKODPPCVARC 600
DB 541 VEECRVLOGLPREVYNARHCLPCHPECOPOGNSVTCFGEADOCVACHYKODPPCVARC 600
QY 601 PSQVPRDLSTYMPIMKFPDEBGACOPCPINCHSCVDLDDKGCPCAPORASPLTISIAYVG 660
DB 601 PSQVPRDLSTYMPIMKFPDEBGACOPCPINCHSCVDLDDKGCPCAPORASPLTISIAYVG 660
QY 661 ILVVVVGAVFGLIKRROQKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
DB 661 ILVVVVGAVFGLIKRROQKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
QY 721 RKVKVLSGAGFTYKGIWIPDGENVKIPVAIKYLRNTSPKANKEIIDEXYVMAGVGSP 780
DB 721 RKVKVLSGAGFTYKGIWIPDGENVKIPVAIKYLRNTSPKANKEIIDEXYVMAGVGSP 780
QY 781 YVSRLLGICLTSTVOLVQLMPYGCGLDHRVNRGRISODLLNMCQIAGKMSYLEBYR 840
DB 781 YVSRLLGICLTSTVOLVQLMPYGCGLDHRVNRGRISODLLNMCQIAGKMSYLEBYR 840
QY 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETEHYADGGKVPDKMMALESILRRFT 900
DB 841 LVHRDLAARNVLYKSPNHVKITDFGLARLLIDETEHYADGGKVPDKMMALESILRRFT 900
QY 901 HOSDWSYGVYVWELMTGAKYDGI PAREIPDLLLEKGBRLPQPICTIIDVYMIWKKWM 960
DB 901 HOSDWSYGVYVWELMTGAKYDGI PAREIPDLLLEKGBRLPQPICTIIDVYMIWKKWM 960
QY 961 IDSECRPRRELVSFBSMARDPORFVYIIONEDLPASPLDSTFYRSLLEDDMDGLVDA 1020
DB 961 IDSECRPRRELVSFBSMARDPORFVYIIONEDLPASPLDSTFYRSLLEDDMDGLVDA 1020
QY 1021 EBYLVPOQGFPCPDAPAGAGVWHRHRSSTRSGGDLTLGLEPSEBEPASPLAPSEG 1080
DB 1021 EBYLVPOQGFPCPDAPAGAGVWHRHRSSTRSGGDLTLGLEPSEBEPASPLAPSEG 1080
QY 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQYSBPTVPLPSETTGVAAPLTCSPQBEV 1140
DB 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQYSBPTVPLPSETTGVAAPLTCSPQBEV 1140
QY 1141 NOPDVRPOPPSPREBGLPAAPAGATLERPKTSLSGKGVVVDVAFCGAVENPEYLLPO 1200
DB 1141 NOPDVRPOPPSPREBGLPAAPAGATLERPKTSLSGKGVVVDVAFCGAVENPEYLLPO 1200
QY 1201 GGAAPQPPAPAFSPAFDNLVYWDODPPERGAPSTFKGTPAENPEYLLGLDVPV 1255
DB 1201 GGAAPQPPAPAFSPAFDNLVYWDODPPERGAPSTFKGTPAENPEYLLGLDVPV 1255

RESULT 2
US-10-770-726-62
; Sequence 62, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE OF INVENTION: CANCERS
; FILE REFERENCE: AM101079 (031896-010000)
; CURRENT APPLICATION NUMBER: US/10/770,726
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 62
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-726-62

Query Match 99.9%; Score 6806; DB 6; Length 1255;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MELAALCRWGLLLALLPPGAASVOYCTGTDMLKRLPASPEYTHLMDLRLLYGGCQVVGNTL 60
DB 1 MELAALCRWGLLLALLPPGAASVOYCTGTDMLKRLPASPEYTHLMDLRLLYGGCQVVGNTL 60
QY 61 ELTYLPTNASLSPLODIOEVQGYVLIANOVROYPLORLRIVRGTOLEFEDNYALAVLDNG 120
DB 61 ELTYLPTNASLSPLODIOEVQGYVLIANOVROYPLORLRIVRGTOLEFEDNYALAVLDNG 120
QY 121 DPLNNTPTVTSASFGRLRELQRLSLTEILKSGVLIORNPOLCYODTILMKDI FHKNNQLA 180
DB 121 DPLNNTPTVTSASFGRLRELQRLSLTEILKSGVLIORNPOLCYODTILMKDI FHKNNQLA 180
QY 181 LTLIDTNRSAKCHSPCKSKRCWGSSEBCCQSLTRVACGCGARCKGRLPTDCCHQC 240
DB 181 LTLIDTNRSAKCHSPCKSKRCWGSSEBCCQSLTRVACGCGARCKGRLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLAFPHNSGICELHCPALVYNTDTPESMPNPEGRTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLAFPHNSGICELHCPALVYNTDTPESMPNPEGRTFGASCVTACP 300
QY 301 YNYLSTDVGSCTIVCPPLHNOEVTADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
DB 301 YNYLSTDVGSCTIVCPPLHNOEVTADGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
QY 361 IOEPAGCKKITGSLAFPLPESFDGPASNTAPLOPQOLOVFTLEITGYLYISAMPDLP 420
DB 361 IOEPAGCKKITGSLAFPLPESFDGPASNTAPLOPQOLOVFTLEITGYLYISAMPDLP 420
QY 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISMGLRSRLRELSGSLALIHNTHTLCFYVTV 480
DB 421 DLSVFQNLQVIRGRILHNGAYSLLTQGLGISMGLRSRLRELSGSLALIHNTHTLCFYVTV 480
QY 481 PMDQLFRNPQALHTANRPEDECVBGLACHOLCARGHCWGPPTQCVCNCSQFLRGQEC 540
DB 481 PMDQLFRNPQALHTANRPEDECVBGLACHOLCARGHCWGPPTQCVCNCSQFLRGQEC 540
QY 541 VEECRVLOGLPREVYNARHCLPCHPECOPOGNSVTCFGEADOCVACHYKODPPCVARC 600
DB 541 VEECRVLOGLPREVYNARHCLPCHPECOPOGNSVTCFGEADOCVACHYKODPPCVARC 600
QY 601 PSQVPRDLSTYMPIMKFPDEBGACOPCPINCHSCVDLDDKGCPCAPORASPLTISIAYVG 660
DB 601 PSQVPRDLSTYMPIMKFPDEBGACOPCPINCHSCVDLDDKGCPCAPORASPLTISIAYVG 660
QY 661 ILVVVVGAVFGLIKRROQKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
DB 661 ILVVVVGAVFGLIKRROQKIRKTYMRRLQETELVEPLTPSGAMPNOAQMRILKETEL 720
QY 721 RKVKVLSGAGFTYKGIWIPDGENVKIPVAIKYLRNTSPKANKEIIDEXYVMAGVGSP 780
DB 721 RKVKVLSGAGFTYKGIWIPDGENVKIPVAIKYLRNTSPKANKEIIDEXYVMAGVGSP 780
QY 781 YVSRLLGICLTSTVOLVQLMPYGCGLDHRVNRGRISODLLNMCQIAGKMSYLEBYR 840
DB 781 YVSRLLGICLTSTVOLVQLMPYGCGLDHRVNRGRISODLLNMCQIAGKMSYLEBYR 840

Qy 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLDIDETEHYADGKVPDKMMALESILRRFT 900
Db 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLDIDETEHYADGKVPDKMMALESILRRFT 900
Qy 901 HOSDWSYGVYVWELMTFGAKPYDGI PARIPDLLEKGERLPQPICTIDVYMIWKKM 960
Db 901 HOSDWSYGVYVWELMTFGAKPYDGI PARIPDLLEKGERLPQPICTIDVYMIWKKM 960
Qy 961 IDSECRPRFRELVSFSSMARDPORFVVIQNEIDLGPASPLDSTFYRSLLDDMDGLVDA 1020
Db 961 IDSECRPRFRELVSFSSMARDPORFVVIQNEIDLGPASPLDSTFYRSLLDDMDGLVDA 1020
Qy 1021 BEYLVPOQGFPCPDPAAGAGVWVHRHRSSTRSGGDLTLGLBSEEBEAPRSLAPSEG 1080
Db 1021 BEYLVPOQGFPCPDPAAGAGVWVHRHRSSTRSGGDLTLGLBSEEBEAPRSLAPSEG 1080
Qy 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPQPEVY 1140
Db 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPQPEVY 1140
Qy 1141 NOPDVRPQPSRREGPLPAARAGATLBERPRTLSPGKGVVDVAFGAVENPEYLTPO 1200
Db 1141 NOPDVRPQPSRREGPLPAARAGATLBERPRTLSPGKGVVDVAFGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPFAFDNLYYWDODPPERGAPPESTFKGTPTAENPEYLGIDVPV 1255
Db 1201 GGAAPQHPHPAPFAFDNLYYWDODPPERGAPPESTFKGTPTAENPEYLGIDVPV 1255

RESULT 3

US-11-113-202-10
; Sequence 10, Application US/11113202
; Publication No. US20050272637A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; APPLICANT: Shantleh, Lara
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1
; FILE REFERENCE: 49321-136
; CURRENT APPLICATION NUMBER: US/11/113,202
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893
; PRIOR FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 10
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-202-10

Query Match 99.9%; Score 6806; DB 7; Length 1255;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1253; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MELAALCGWGLIALLLPGCAASTQVCTGDMKLRLPASBETHLDMLRHLYGCGQVQGANL 60
Db 1 MELAALCGWGLIALLLPGCAASTQVCTGDMKLRLPASBETHLDMLRHLYGCGQVQGANL 60
Qy 61 ELTYLPTNASLFLDDIOEVQGVYLIANOVQVLOLRIRYRGQLPEPDNALAVLNG 120
Db 61 ELTYLPTNASLFLDDIOEVQGVYLIANOVQVLOLRIRYRGQLPEPDNALAVLNG 120
Qy 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHNQOLA 180
Db 121 DPLNNTTPTVGTASPGGLRELQRLSLTEILKGGVLIQRNPOLCYODTILMKDIFHNQOLA 180
Qy 181 LTLITNNSRACHPCSPMKGSRGSESEDCOSILRTYVACAGCARCKGRLPTDCHBQC 240
Db 181 LTLITNNSRACHPCSPMKGSRGSESEDCOSILRTYVACAGCARCKGRLPTDCHBQC 240

Qy 241 AAGCTGPKHSDCLACIAHFNHSGICELACPALVTYNTDTPESMNPBEGRYTFGASCVTACP 300
Db 241 AAGCTGPKHSDCLACIAHFNHSGICELACPALVTYNTDTPESMNPBEGRYTFGASCVTACP 300
Qy 301 YNYLSTDVSGCTLVCPPLNQBVTAEADGTORCEKSKPCARVYCYGIMELREVRVATSAN 360
Db 301 YNYLSTDVSGCTLVCPPLNQBVTAEADGTORCEKSKPCARVYCYGIMELREVRVATSAN 360
Qy 361 IOEFAGCKKIKTSIALFLPESFPDGPASNTAPLOPQLOVFEETLEBITGLYLSAMPDSLAP 420
Db 361 IOEFAGCKKIKTSIALFLPESFPDGPASNTAPLOPQLOVFEETLEBITGLYLSAMPDSLAP 420
Qy 421 DLVSQNLQVIRGRILIHNGAYSILTQGLGISWLAGRSLRELSGGALIHNNTHLCFVHTV 480
Db 421 DLVSQNLQVIRGRILIHNGAYSILTQGLGISWLAGRSLRELSGGALIHNNTHLCFVHTV 480
Qy 481 PWDQLFRPHQALLHTANRPEDECVGEGIAHQLCARGHCWGPPTQCVNCSQFLRGQBC 540
Db 481 PWDQLFRPHQALLHTANRPEDECVGEGIAHQLCARGHCWGPPTQCVNCSQFLRGQBC 540
Qy 541 VEECRVLOGLPREYNABHCLPCHRECOPONGSVTCFGEADQCVACHYKDPFCVARC 600
Db 541 VEECRVLOGLPREYNABHCLPCHRECOPONGSVTCFGEADQCVACHYKDPFCVARC 600
Qy 601 PSQVPRDLSTYMPITWTFPDEBGAQCPPICTHSCVDLDDKGCAPQORASPLTSIISAVVG 660
Db 601 PSQVPRDLSTYMPITWTFPDEBGAQCPPICTHSCVDLDDKGCAPQORASPLTSIISAVVG 660
Qy 661 ILVVVIGVFGIILIKRQOKIRKTYMRLLQETELVEBLTPSGAMPNOQMRILKETEL 720
Db 661 ILVVVIGVFGIILIKRQOKIRKTYMRLLQETELVEBLTPSGAMPNOQMRILKETEL 720
Qy 721 RKVKVLGSGAGCTYVKGIMIPDGENVKI PVAILKVLRNTSPANKETILDEAYVMGVGSP 780
Db 721 RKVKVLGSGAGCTYVKGIMIPDGENVKI PVAILKVLRNTSPANKETILDEAYVMGVGSP 780
Qy 781 YVSRLLGICLSTVQVLTQVLMPEYGCILDHVENRGRLSODLIANCMQIAKMSYLEDR 840
Db 781 YVSRLLGICLSTVQVLTQVLMPEYGCILDHVENRGRLSODLIANCMQIAKMSYLEDR 840
Qy 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLDIDETEHYADGKVPDKMMALESILRRFT 900
Db 841 LVHRDLAARNVLYKSPNHVKITDPGLARLLDIDETEHYADGKVPDKMMALESILRRFT 900
Qy 901 HOSDWSYGVYVWELMTFGAKPYDGI PARIPDLLEKGERLPQPICTIDVYMIWKKM 960
Db 901 HOSDWSYGVYVWELMTFGAKPYDGI PARIPDLLEKGERLPQPICTIDVYMIWKKM 960
Qy 961 IDSECRPRFRELVSFSSMARDPORFVVIQNEIDLGPASPLDSTFYRSLLDDMDGLVDA 1020
Db 961 IDSECRPRFRELVSFSSMARDPORFVVIQNEIDLGPASPLDSTFYRSLLDDMDGLVDA 1020
Qy 1021 BEYLVPOQGFPCPDPAAGAGVWVHRHRSSTRSGGDLTLGLBSEEBEAPRSLAPSEG 1080
Db 1021 BEYLVPOQGFPCPDPAAGAGVWVHRHRSSTRSGGDLTLGLBSEEBEAPRSLAPSEG 1080
Qy 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPQPEVY 1140
Db 1081 AGSDVFDGDLGMAKGLQSLPTHDPSPLOQYSEDPVPLPSETDGYVAPLTCSPQPEVY 1140
Qy 1141 NOPDVRPQPSRREGPLPAARAGATLBERPRTLSPGKGVVDVAFGAVENPEYLTPO 1200
Db 1141 NOPDVRPQPSRREGPLPAARAGATLBERPRTLSPGKGVVDVAFGAVENPEYLTPO 1200
Qy 1201 GGAAPQHPHPAPFAFDNLYYWDODPPERGAPPESTFKGTPTAENPEYLGIDVPV 1255
Db 1201 GGAAPQHPHPAPFAFDNLYYWDODPPERGAPPESTFKGTPTAENPEYLGIDVPV 1255

RESULT 4
US-11-113-202-6
; Sequence 6, Application US/11113202
; Publication No. US20050272637A1


```

Db 74 YDLSEFKTIOEVAGVYLIALNTERIPIENLOIRGNMYENSVALAVLSND----- 126
Qy 129 VTGASPGGIRELQSLSTEILKGVLIQBNPOLCYODTILMKDIFHKNOQLATLIDTR 188
Db 127 ---ANKTELKEIPMANLOEILHGAIVRFSNNPALCVESIQMRDIYSSDPLSMMSMDPQNH 183
Qy 189 SRACHPCSPMCKSGHSCWGESSEDCQSLTRTVCAAGCA-RCKGRLPTDCCHQCAAGCTGP 247
Db 184 LGSQCKDPSCPNCSGWCAGEBNCQKLTIKICAOCCSGCRGKSPSDCHNCAAGCTGP 243
Qy 248 KHSDDLACHFNHSGICEILHCPALVTYNTDTFESMPNBEGRTRTBASCTYACPYNYLSTD 307
Db 244 RESDCLVRKRDEATCKDTCPPMLYNPPTYQMDVNPBGKXSFQATCYKCKPRYVYVD 303
Qy 308 VGSCTLVCPRLHNOETAEEDGTORCEKSGKPCARVCYGLMEHLREVRATVSANIOEPAGC 367
Db 304 HGSCYRACGADSYEN-EEEDGVAKCKCKGCPKVCNGIGIGFCKDLSINAINIHFKNC 362
Qy 368 KKFISLAFLEPESPDGDPASNTAPLQBPOLQVFTLEETGYLYISAMPDSLPLDSVFQ 427
Db 363 TSIISGDLHLPLVAFRGDSFTHTPPLDPOBLDKTKVKEITGFLLIQAMPENRTDLHAFEN 422
Qy 428 LQVIRGRILHNGAYSILTQIGISWLGRLSRLSELSGLALIHNTHLCPVHTVPDQLFR 487
Db 423 LEIRGRTRKHQGFSLAVVSLNITSLGRSLKEISDGVYISGNKLCYANTINMKKLF 482
Qy 488 NPHALHTANRPECEVGEGLACHQLCARHCGWGPPTOCVNSQFLRGCEVCECHYL 547
Db 483 TSGQCTKILISNNGENSCATGVCALCSPEGCWGPBPDVCSKNVSRGECVCKWCL 542
Qy 548 QGLPREYVNAHCLPCHPECPONGSVTCFGEADQCAVACHKDPFCVACSPGVKPD 607
Db 543 EGBPREFVENSBCIQCHPECLPQAMNITCTGRGPNCIQCAHYIDGPHCVKTCRPAWGME 602
Qy 608 LSWPIKFRPEBEGACQPCPINCHSCVLDLDCGCPABORASPLTSIISAVG---ILIV 664
Db 603 NNLT-VMKADAGHCHLCHPCTGCTGPGLEGCPYKPKP--SIATGMGALLLILIV 659
Qy 665 VLVGVVPGIILIKRROKIRKYTMRLLOETELVEFLTPSGAMPNOAKRIIKETELRVK 724
Db 660 VALGIG---LPMRRHIVKRLRLLOERELVEFLTPSGEAPNALIKLETFPKIK 716
Qy 725 VLGSGAFSTVYKGIWIPGENVKIPVAIKVLRENTSPYANKELIDEAVYMAVGSPVYSR 784
Db 717 VLGSGAFSTVYKGIWIPGENVKIPVAIKVLRENTSPYANKELIDEAVYMAVSPVYSR 776
Qy 785 LLGICLSTVOLVLOLMPYGCILDHVRENRGLSGQDILNMCQIAKMSYLEDVRLVHR 844
Db 777 LIGICLSTVOLVLOLMPYGCILDHVRENRGLSGQDILNMCQIAKMSYLEDVRLVHR 836
Qy 845 DLARNVTVKSPNHVKITDPGLARLDIDETEVHADGGKVPYKMMALSSILRRFTHSD 904
Db 837 DLARNVTVKSPNHVKITDPGLARLDIDETEVHADGGKVPYKMMALSSILRRFTHSD 896
Qy 905 VMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGRLLPOPICITIDVYMIWYKMMIOSE 964
Db 897 VMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGRLLPOPICITIDVYMIWYKMMIOSE 956
Qy 965 CRPRRELIVSEFSRMAPDPQRFVVIQ-NEDUGPASPLDSTYRSILLEDMDGDLVDAEY 1023
Db 957 SRPKRELIIFFSKARDPORYLVIQDGBRMHLPSPTDSNFYRALMDDEMDVDVDAEY 1016
Qy 1024 LVPOGFGFCPRPARAGGMVHHRRSSSTRSGGDLTGLBFSSEEARSPLABSEGGS 1083
Db 1017 LVPOGFGFCPRPARAGGMVHHRRSSSTRSGGDLTGLBFSSEEARSPLABSEGGS 1042
Qy 1084 DVFPDDLGMGAAGLQSLPTHDPSPLQRYSEDPVPLPSET--DGYVAPLTCSPQPEYVN 1141
Db 1043 N-INSTVACIDRNGQSCPIKEDSFLQRYSSDPTCALIEDSIDDFL-----PPEYIN 1094
Qy 1142 QPDVAPPPSPREGVLPAARPAATL-----ERPYYLSPGKNGVVKDFVAFAGAVENDEY 1196
Db 1095 Q-----SVP-KRPAQSVONPVYHNOPLNPAASRDPHYD--PHSTAVGNDEY 1138

```

```

Qy 1197 L-TPGGAARQPHRPPAPAFDNLXYWD-----DP-----PERGAPSTFKG 1239
Db 1139 LNTVQ-----PTCVNSTFDSPAHNAQKSHQISLNDPVOQDFFPKKAPNGIFKG 1189
Qy 1240 TPTAENPEYL 1249
Db 1190 S-TAENAEYL 1198

RESULT 6
US-11-113-202-16
; Sequence 16, Application US/11113202
; Publication No. US20050272637A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1
; FILE REFERENCE: 49321-136
; CURRENT APPLICATION NUMBER: US/11/113,202
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893
; PRIOR FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16
; LENGTH: 1308
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-202-16

Query Match 44.1%; Score 3004.5; DB 7; Length 1308;
Best Local Similarity 45.6%; Pred. No. 2.2e-183;
Matches 614; Conservative 183; Mismatches 375; Indels 175; Gaps 27;

9 MGLLALLRPPGAA-----STOYCTGTDMKRLPASPETHLDMRLHLYOGCOVVQNLIELTY 64
8 WWSVSLVAAGTVPDSQSVCAGTENKLSLSLEQGYRLRKTYENCYVMGHEITS 67
65 LPTNASFLDIOEVQGVLIANHOVQVPLQRLIRVGTQLEFEDNYALAVLDNGDPLN 124
68 IEHRNDSFLKSVREVTGYVALNQFRYLPLENRIIRGKLYEDRVYALAIFLNYRKDG 127
125 NTPYTGASPGGLBELQSLSTEILKGVLIQBNPOLCYODTILMKDIFHKNOQLATLII 184
128 NF-----GLOBELGKMLTEILMGVYVDQNFCLCYADTINMODIVRNPWPSNLTIV 178
185 DTNRSRACHPCSPMCKSGHSCWGESSEDCQSLTRTVCAAGC-ARCKGPLPTDCCHQCAAG 243
179 STNSSGCGRCHKSGTG-RCKGPTENHCQTLIRTVCAQCGRCYGPVYSDCHRECKAG 237
244 CTGPGISDCLALFHFNHSGICEILHCPALVTYNTDTFESMPNBEGRTRTBASCTYACPYNY 303
238 CSGPRDIDCFQAMNPNDSGACVTCQPFQFVYNNPTTFOLEHFNPAKYTGACVCKCPHNF 297
304 LSTDVSGCTLVCPRLHNOETAEEDGTORCEKSGKPCARVCYGLMEHLREVRATVSANIOE 363
298 V-VDSSSCVRAQSPSSKMEV-BENGIKMKCKPCTDICPKKADIDIGTSLMSADTAVSSNIDK 355
364 FAGCKIFGSLAFLEPESPDGDPASNTAPLQBPOLQVFTLEETGYLYISAMPDSLPLDS 423
356 FINCKINGNLIFLVGTGHPYNAIEAIDDEKLVNFTVEITGFLIQTSMPPMTDPS 415
424 VFQNLQVIRGRILHNGAYSILTQIGISWLGRLSRLSELSGLALIHNTHLCPVHTVPD 483
416 VFSNLVTGKGVLSGLSLILKQGGITSLFQSLKEISAANIYITDSNICYHTTIWT 475
484 QLRNPHQALHTANRPECEVGEGLACHQLCARHCGWGPPTOCVNSQFLRGCEVCEY 543
476 TLFPSTINQRIIVIRDNRKAKENCTAGAMVCNHLSSSDGCGWGPBPDCLSRFRSRRGICIES 535

```

```

QY 544 CRVLQGLPREYVNAHCLPCHPEQOP-ONGSVTCTGPERADCVACAHKQOPRFVACPS 602
QY 536 CNLVDGEEREPENGSI CYECBPQCEKMEGDLTGHPEPNDCTYKSHPKQSPNCVCEKCPD 595
QY 603 GVKPRLSTWPLTWKPFDEEGACORPEINCTHSCVDLDDKGC-----PAEQASPL 651
QY 556 GLQGANSE--LFPKVPDPRECHPCHPNTCTOGCNGPTSHDCIYYWYTGSHTLPGAR---- 649
QY 652 TSIIISAVV--GILLVVLGVVFGIILIKRQOKIRKYTRRLLOETELVEPLTPSGAMPNQ 709
QY 650 TPLIAGYIGGLFIVILVIGLTFEAVVYRRKSIK-KKRALRREL-ELELVEPLTPSGTAPNQ 707
QY 710 AQMRILKTELAKXVVLGSSGAFGTYWKIWI PDGNNYKI PAIKVLRBNTPSPKANKLELD 769
QY 708 AQRLIKTELKRVVVLGSSGAFGVYKGIWPEGTIVKI PAIKILNETHTPKANVEFMD 767
QY 770 EAYVWAGSPVYSRLGLICTSTFVOLTQTMPEYSCLDHYRENRGRIGSODLLNMCQOI 829
QY 768 EALIMASDHRLVVLGVLGCLSPITQVLTQMLPHCCLLEYHKKNDNISGDLNMCVQI 827
QY 830 AKGMSYLEDVRLVHRDLAARNVLKSPNHVKITPBGRLALIDIDETEHADGKVPYIKM 889
QY 828 AKGMVYLEBERLVHRDLAARNVLKSPNHVKITPBGRLALIDIDETEHADGKVPYIKM 887
QY 890 ALESLIRBRPHNOSVMSYGVTWMLMFFGAKPDGIPAREIPDLLEKGEHLPOPICTI 949
QY 888 ALBCTHYKFKHQSDVMSYGVTIMLMFFGKPDGILTRKELPDLLEKGEHLPOPICTI 947
QY 950 DVMYIMVXCMWIDSECRPFRELYSEFSRMARDPQRFVYIIONED-LGPASPLDSTFYRSL 1008
QY 948 DVMYVWVXCMWIDADSRPKFKELAEFSRMARDPQRFVYIIGDDRMKLPSPNDSKFPONT 1007
QY 1009 LEDDMGVLVDAAEVLVLVQQGFPCDPAFGAGMWNHHRSSSTRSGGDLTLGLESEB 1068
QY 1008 LDEELEDMDMAEEYLV- QAFNIPRP-----LYTSRARLDSNRS-----EIGHSPRA 1055
QY 1069 EAPRS-----PLAP-SEGASDPVFGDGLGMAAKQLQ 1099
QY 1056 YTPWNGNQVYRBDGFAAEQGVSVYRPAFTSTIRPAPAAQGTATIRPDSCTNGLTKRVP 1115
QY 1100 SLPTHPSPRLQRYSEDPTVPPLPS-----ETDGYAULCTSPQPEYVNOQDVBPQRPSP 1152
QY 1116 APHVQEDSSTQGVYADPTVPAPERSPRGELDBEGYMTBMRKPKQEXYLNPIVE----- 1167
QY 1153 REGPAPARPGATILERKTLSPGKNGYVVKVPAFGAVENPEYTLTPGGAARQHPHPPA 1212
QY 1168 -ENPFVSR-----KNGLQ-----ALDNPYHNASNG-----PEKA 1198
QY 1213 -----FSPAFDMLYVWDOPPERGA--PPST 1236
QY 1159 EDEYVNEPLVNTFANTLGKAEYLNKNLILMSPEKAKKAFDNPDTVWNHSLPFRSLQHPDY 1258
QY 1237 FKGPRT-----AENPEYL 1249
QY 1259 LQEVSTKYFYKQNGRIRPIVAENPEYL 1265

RESULT 7
US-11-113-202-8
; Sequence 8, Application US/11113202
; Publication No. US20050272637A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; APPLICANT: Shantele, Lora
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1
; TITLE OF INVENTION: RECEPTOR AND ERBB RECEPTORS
; FILE REFERENCE: 49321-136
; CURRENT APPLICATION NUMBER: US/11/113,202
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893

```

```

; PRIOR FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3.1
; SEQ ID NO 8
; LENGTH: 943
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-202-8

```

Query Match	37.4%;	Score 2545.5;	DB 7;	Length 943;
Best Local Similarity	51.3%;	Pred. No. 2e-154;		
Matches 510;	Conservative 136;	Mismatches 245;	Indels 103;	Gaps 19;

[illegible]

```

Qy 1173 LSPGKGVKDVAFAGAVENPEYL-TPGGAAPQHPPEAPSPAFDNLVYWDQ----- 1225
Db 850 PAPSPDHPHQD--PSTAVNGPELINTVQ-----PTCVNSTFTDSPAHAQAQSHOI 898
Qy 1226 ---DP-----PERGAPSTFKGTPTAENPEYL 1249
Db 899 SLNDNDYQODFPPEAKPWFIFKGS-TAENAEYL 931

RESULT 8
US-10-770-726-63
; Sequence 63, Application US/10770726
; Publication No. US2005026409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; TITLE OF INVENTION: CANCERS
; FILE REFERENCE: AM101079 (031896-010000)
; CURRENT FILING DATE: 2004-02-04
; NUMBER OF SEQ ID NOS: 48640
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 63
; LENGTH: 1342
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-726-63

Query Match 35.8%; Score 2441.5; DB 6; Length 1342;
Best Local Similarity 41.0%; Pred. No. 1.3e-147;
Matches 539; Conservative 182; Mismatches 457; Indels 137; Gaps 34;

Qy 10 GLLALLPFGAA--STQVCTGTDMLRLPASPETHLDMRLHYQSGOVVQGNLELYLPT 67
Db 11 GLLPFLAGSEVGNQAVCGPTGLNGSLVTDAGENQVQLYKLYRCEVWGNLELYLNGH 70
Qy 68 NASLFLQDIOVQGVYLIANHQVQVPLQRLRYRGVQLFEDNALVALNDGDLNNT 127
Db 71 NADLFLQWIRVTVYLVANNEFTLPLNLRVVRGVQVYDGKALIFW-----LNVLT 125
Qy 128 PVTGASPGGLRELQRLSTEILKGGVLIQRIQNPQLCYQDTILMKDIFHKNNQALATLIDTN 187
Db 126 -----NSSHALRQLRLQTLTEILSGGYIEKNDKLCMDITIDPRDLYRDRD---AETVYKD 178
Qy 188 RSRACHPCSPMCKSGRCSWGESESDCSLTRIVCAGGC-ARCKGRLPTDCHEBOCAAGCTG 246
Db 179 NGRSCRPCHGVCKG-RCMGPGSEDCQTLTKTICAPQCNCHGFRPNPQCHDECAAGGCG 237
Qy 247 PKHSQCLACLHFNHSGICELHCPALVYTYTDTFESMNPNEGTYTGASCVTACPRNYLST 306
Db 238 PDDTCFACRHFNDGACVPRCPPLVYVKKLTFQLEPNHTKYQVGVCVASCPHFV-V 296
Qy 307 DVGSCVLCPLHNOEVTADGTFORCEKSKPCARVCYGLGMEHLBVRVATSANIOEPAG 366
Db 297 DQTSQVRAQCPPEKMEVD-KGGLKMECPCCGLCPKACBGTGSG-SRFQVDDSSNIDGFVN 353
Qy 367 CKKIFGSLAFLPESFDGPASNTAPLOPQLOVFEETLEETGYLYISAMPDSLPLDSVFO 426
Db 354 CTKILGNDFLITGLNGDPMHKIRPALDPEKLNVPFRVREITGYLNIQSPPHMNFVSYS 413
Qy 427 NLQVIRGRILNNGAYS-LTLQGLISWGLRLSLRELGSGLALIHNTLHCFHTYVPMQL 485
Db 414 NLTYTGSGSLYVRGSLIMKNLANTSLGFRSLKEISARIVYISANROLCYHRSLSNMVKV 473
Qy 486 FNPQALHTTA-NRPDECEVGEGLACHQICARGHGWGPPTQCVNCSQFLRGCEVREC 544
Db 474 LRGPTEBRDLIDIGNPRRDCVAGKVCDDLCSGGCKMGSPQCLSCRYSGVGYCTHC 533
Qy 545 RVLQGLPREYVNAHCLPCHPECOPONGSVTCFGBPADQCVACAHYKDPFCFVACPSGV 604
Db 534 NFLNGEPREFAHBAECFSCHPPECOPMGATCNGSGSDTCAQCAHFRDQPHCVSSCPHGV 593

```

```

Qy 605 KEDLSYMPWKPEDEGACOPCPINCTHSQVLDLDDKCPAEORA-----SPLTSIISAVG 660
Db 594 LG-AKGPYKYPDVQNECRPCHECTOCCKPBLQDCLGQVLVLIGKTHLMALTVING 651
Qy 661 ILVLVLGVFGILIKRQOKIR-KYTMRLLOETELVEPLTPSGAMPNOAMRLIKETE 719
Db 652 --LVVIFMWLGGTFLYWGRGRIQNRAMRYLERGESIEPLDPS-EKANKVLARIFKEYE 708
Qy 720 LKXVVLSSGAGCTYTKGIWIPDGNMVKTPVALIKYLRNTSPKAKEILDEAVYVAGVS 779
Db 709 LRKLVLSSGVGVYHKGWIPESGISIKI PVCIKYIEDSGQSFQAVLDHMLAGSLDH 768
Qy 780 PVSRLGLICLTSTQVLTQVLMPCYLGDHVENRGLASQDLMMQOIAKMGYLEDY 839
Db 769 AHVVLGLICSSSIQVLTQVYPLGSLDHYRQHGLAGPQLLMMQVQIAKMGVYLBH 828
Qy 840 RLVRDLAARNVLKSPNHVKITDFGLARLDIDETEHADGKVPIKMALESILRRF 899
Db 829 GNVHNNLAARNVLKSPQGVADFGVADLLRPDDKQLLYSEAKTPIKMMALSIHPKY 888
Qy 900 THQSDVWSYGVTVWELMTFGAKPYDGIAPAKEIPDLLEKGERLPPPICTIDVMMYK 959
Db 889 THQSDVWSYGVTVWELMTFGAEPYAGRLAEVPLLEKGERLAQPOICTIDVMMYK 948
Qy 960 MIDSECRPRELVEEFSMARDDPORFVYIQNEIDGPA---SPLSTFYRSLIEDDMGD 1016
Db 949 MIDENIRPTFKELANEFTRMARDPRRYLYIKES-GPGIAPRPEHGLTNKLEVELEP 1007
Qy 1017 LVDAEYLVPOQGFPCPPAPGA---GGMVHRRHSSTRS-----GGDL 1059
Db 1008 ELDDLDELAEEDNLATTLTGLSLPLVGTILNRPGSGLSPSSGYPMQNGILGSGCQ 1067
Qy 1060 TGLPSESEEARSLA-----PSEGASDVPFGDLGMKAALQD-----LPHDS 1107
Db 1068 ESAVSGSSBRCP-RVSLHPMPRGCLASESSGHV-TSEEARLOKRVSMCRSSRSR 1125
Qy 1108 P-----LQRYSDPTVPL-----PSRTGYVAPLITSPQPEVNOQVPRPQ-PPSP 1152
Db 1126 PRGDSAYHSQHSLLTPVTPPLSPGLSEEDVNGVY-----MPDHLKGTSPS 1172
Qy 1153 REGPLPAARPGATLERPRLTSPGKGVV-----KDYFAFGAVENPEYLTPOGGAPO 1206
Db 1173 REGTSS-----SVGLSVLTDEEBED-----EEVEYMMRRRRHSP- 1208
Qy 1207 PHPPEAPSPAFDNLVYWD-----QDPERGAPSTFKGTPTAENPEYL 1249
Db 1209 PHPPESSLBELGYEYMDVGSLSASLSTQSCPLHPVPVIMPTAGTTDDEYEM 1263

RESULT 9
US-11-113-202-12
; Sequence 12, Application US/11113202
; Publication No. US20050272637A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; APPLICANT: Shantel, Lara
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1
; TITLE OF INVENTION: RECEPTOR AND ERBB RECEPTORS
; FILE REFERENCE: 49321-136
; CURRENT APPLICATION NUMBER: US/11/113,202
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12
; LENGTH: 1342
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-202-12

```

Query Match 35.8%; Score 2441.5; DB 7; Length 1342;
 Best Local Similarity 41.0%; Pred. No. 1.3e-147;
 Matches 539; Conservative 182; Mismatches 457; Indels 137; Gaps 34;

```

10 GLLALLPPGAA--STGYCTGTMKRLPABPETHLDMRLHYGQCQVYQGNLEITYLPT 67
11 GLFLSLARGSEVGNQAVCPGTLNGLSVTGDENQYQTLYLKERCEVVMGNLEIVLIGH 70
68 NASISFLDIOEVGYVLIANQVAVPLQRLRIYRGQLEPENVALLVNDGPIANNTT 127
71 NADISFLQWIREVGYVLIANNEFSTLPPLNLRVVRGTQYVDGKFAIFVM-----LNNNT 125
128 PVTGASPGLEBLQRLSLEILKGGVLIQRPOLCYODTILMKDIFHNQNLALTLIDTN 187
126 ----NSSHALRQLRLTQTLTEILSGGYIEKNDKLCMDTIMRDIVRDRD---AEIVKD 178
188 RSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGC--ARCKGRLPTDCCHGQCAAGCTG 246
179 NGRSCPCHGVCKG--RCWGPGESEDCQTLTKTICAPQCNHCGFPNPNCCCHDECAGGCSG 237
247 PKHSICLACLFHNSGICELHCPALVTYNTDFESMPREGRYTFGASCYTAAPYNYLST 306
238 PDDTDCFACRPHNDGACVPRCPQPLVYNNKLTFLQLEPNPHTKYQYGVCAVASCPIHFV--V 296
307 DVGSCCTVCPHNOEVTAEADGTQRCCKSKPCARVCYGLGMEHLREVAATVSANIOEPAG 366
297 DQTSCTVACRPDKMEVD--KNGLKMEPEPGGLCPKACBGTSG--SRFTVDSNIDGVN 353
367 CKTIIFGSIAPLPESFDGPASNTAPLQEPQLOVFETLEITGYLYISAMPDLSPLSFQ 426
354 CTKIIGNDFLITGLNGDPWHNKIPALDEKLNVEFTVEITGYLNIQSPMPHNFVSFS 413
427 NLQVRGELIINGAAS--LTLOGLGSLWGLSLRSLRELSGLALIHNNHLCEVHTYPMQOL 485
414 NLTTIGSLNKGSLIMKMLNVTSLGPNLSLKISGRYISANROLCHHSIMTKV 473
486 FRNPQALLHTA--NRPEDEVCEGLACHQLCARHGMCPGPTQCVNCSQFLRGQCEVCEC 544
474 LRGPTEERLDIKHNRPRRDCAVAEGKCDPLCSGGCKWGRGQGLSCNRYRGCVYHNC 533
545 RVLQGLPREYVNAHCLPCHRECPQONGSVTCFGEADQCAVAHAKDPRECVARCPSEGV 604
534 NFLNGEPREFAHEACFSCHPECCQPMGGATCNCGSDYCAQCAHFRDGPCHVSCSPHGV 593
605 KPDLSYMIKPRDEBGAQCPPTICNCTHSQVDLDKGPABQRA-----SPLTISIAYVG 660
594 LG--AKGPIYKRPDVQNECRPCHECTGCGKGPBLQDCLGQTLVLIGKTHLTMALTVIAG 651
661 ILLVVLGCVGILIKRRQOKR--KYTWRRLLQETELVEPLTPSGAMPNOAQMRILKETE 719
652 --LVVIFPMGLSTPLYMGRRIQNRAMRARIYERGESIEPLDPS--EKANKYLARIFKETE 708
720 LRKVKVLGSGAFGYVYKGIWIPDGENVKIPVAIKYLRENTSFPKANKELIDEAYVMAGVGS 779
709 LRKLKVLGSGVGYVHKGVMIPEGESIKIPVCIKYIEKSGROSFOAVTDHMLAGSIDH 768
780 PYSRLILGICLSTQVLTQLMRPGVCLDHYRENNGRIGSODLNMCMQAIKMGSYLBDV 839
769 AHTVVLGICPRESSIQLVTVQYPLGSLLDHYRQNHGALGAPOLLNMGVQAIKMGYLBEN 828
840 RLVRHDLAARVNLVYSPNHVKITDFGLARLLDIDETEYHADGKVPRIKMMALLESILRRP 899
829 GMVHRHNLAAARVNLKSPGOVAVDGVADLPRDDKOLLYSEAKPIRIMMALLESIHPEKY 888
900 THQSDVMSGYVTWELMTFGAKPYDGIPAREIPDLLEKGERLPORPCTITIVYMTVKW 959
889 THQSDVMSGYVTWELMTFGAEPYAGLRILAEVPDLLEKGERLAPQICTIVYMMYVKKM 948
960 MIDSCRRPRELVESEFARMADPGRFVIVQINEDGPA---SPLDSTYRSLLEDDMDGD 1016
949 MIDENIRPTFKELANEFTRMADPPRYLVTKRES--GPGIAPGPEPHGLTNKKLBEVELEP 1007

```

```

1017 LVDAEYLVPOQGFCEPDPAFGA---GMYNHRHRSSTRS-----GGDL 1059
1008 ELDDLDELEEDMLATTLGASLSTLPVGTILNRRPGCSQLSPSSGYPMNQNLGSGSCQ 1067
1060 TLGLEPDEEERPRPLA-----PSGAGSDVFDGLMGAAKGLDS-----LPTHDS 1107
1068 ESAVSGSERCP--PVSILHPNRCGLASESEGHV--TQSELELDEKVMCSRSRSRSPR 1125
1108 P-----LQRY--EDTPVPL-----PSETDGYAALPSCPOPEYVNPDPVRQ--PSP 1152
1126 PRGSAHSGHSHSLTFTVPLSPRGLSEEDYNGTV-----MPDTHLKGTSS 1172
1153 REGPLPAARPAAGATLERPKLSPKNGV-----KQVFAGAVENPEYLTPOGGAAPQ 1206
1173 REGTLS-----SVGLSSVLTGEDEED-----EEYENNRRRRHSF-- 1208
1207 PHPPAPSPAFDNLVYMD-----ODPERGAPSPFTKGTPTAENPEYL 1249
1209 PHPPPSLLEELGYEYMDVGSLSLSTQSCPLHPVPLMPTAGTTPDEDEYEM 1263

```

```

RESULT 10
US-11-113-202-14
; Sequence 14, Application US/11113202
; Publication No. US20050272637A1
; GENERAL INFORMATION:
; APPLICANT: Shmikh, Lara
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1
; TITLE OF INVENTION: RECEPTOR AND ERBB RECEPTORS
; FILE REFERENCE: 49321-136
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893
; PRIOR FILING DATE: 2004-04-22
; SOFTWARE: PatentIn version 3.3
; LENGTH: 1342
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-202-14

```

Query Match 35.8%; Score 2441.5; DB 7; Length 1342;
 Best Local Similarity 41.0%; Pred. No. 1.3e-147;
 Matches 539; Conservative 182; Mismatches 457; Indels 137; Gaps 34;

```

10 GLLALLPPGAA--STGYCTGTMKRLPABPETHLDMRLHYGQCQVYQGNLEITYLPT 67
11 GLFLSLARGSEVGNQAVCPGTLNGLSVTGDENQYQTLYLKERCEVVMGNLEIVLIGH 70
68 NASISFLDIOEVGYVLIANQVAVPLQRLRIYRGQLEPENVALLVNDGPIANNTT 127
71 NADISFLQWIREVGYVLIANNEFSTLPPLNLRVVRGTQYVDGKFAIFVM-----LNNNT 125
128 PVTGASPGLEBLQRLSLEILKGGVLIQRPOLCYODTILMKDIFHNQNLALTLIDTN 187
126 ----NSSHALRQLRLTQTLTEILSGGYIEKNDKLCMDTIMRDIVRDRD---AEIVKD 178
188 RSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGC--ARCKGRLPTDCCHGQCAAGCTG 246
179 NGRSCPCHGVCKG--RCWGPGESEDCQTLTKTICAPQCNHCGFPNPNCCCHDECAGGCSG 237
247 PKHSICLACLFHNSGICELHCPALVTYNTDFESMPREGRYTFGASCYTAAPYNYLST 306
238 PDDTDCFACRPHNDGACVPRCPQPLVYNNKLTFLQLEPNPHTKYQYGVCAVASCPIHFV--V 296
307 DVGSCCTVCPHNOEVTAEADGTQRCCKSKPCARVCYGLGMEHLREVAATVSANIOEPAG 366
297 DQTSCTVACRPDKMEVD--KNGLKMEPEPGGLCPKACBGTSG--SRFTVDSNIDGVN 353

```

```

Qy 367 CCKIFGSLAFLBESPDGASNTAPLOPELOVFEETLEITGVLISAMPDLSLDSVQ 426
Db 354 CCKIGLNDLFLTGLNGDPMHKKIPLADPEKLVNFRVIRITGLINQSPRHMHNSVYS 413
Qy 427 NLQVIRGRILHNGANS-LTLQIGISWLGIRSLREISGLALIHNTHLCPVHTVMDOL 485
Db 414 NLTTIGRSLYVRGFSLLIMKQNLMTVSLGFRSLKEISAGRIVISANROLCYHNSLMTKY 473
Qy 486 FNNPQALLHTA-NREDECVGEGLACHOLCARGHCMGWPQCVNCSQFLRGQEVRESC 544
Db 474 LRGPTBEERIDIKNNRRRCVAGKVCDDLCSGGCMGPGQCSGRNYSRKGVCVTHC 533
Qy 545 RVLQGLPREYVNAARHCLPCHPECOPOGSGVTCGPEPADOCVA CANHYKDPFVCARCPSCV 604
Db 534 NFLNGEPRFPAHABEFCSCHECOPEGRATOCGSSDVCACAHNRDOPHCSSCPHEV 593
Qy 605 KQDLSYMPKRPDEEGACQPCPINCTHSCVDLDDKCPAEORA---SPLTSIISAVVG 660
Db 594 LG--AKGPYKYPDVQNECRPCHEMCTQCKGPELDCLGQTLVLIGKTHLTMALTVIAG 651
Qy 661 ILLVVLGVVPGILLIRROOKIR-KYTMRLLOETELVPLTPSGAMPNOAMRLIKETE 719
Db 652 --LVVIFMVLGGFTLYWRGRRIQNKRAMRYLERGESIEPLDPS-BKANKVLARLPKETE 708
Qy 720 LARKVLSGSAFGTVYKGIWIPDGBENVKI PVAIKVLRBENTSPYANKRIIDEAYVMAVGVS 779
Db 709 LAKVLKSGVGTYNKGVWIPBESIKIPVCIKVIEDSGSGOSPOAVTDHMLAIGSLH 768
Qy 780 PYVSRLLGICLTSTVQVLTQMLPYGCLLDHVENRGRIGSODLLMCMQIAKMSYLEDV 839
Db 769 AHIVRLGLCPGSSILQVLTQVLPGLSLDHYVROHGRALPQLLLMNGVLIAGMYLEEH 828
Qy 840 RLVNHDLAARNLVLSRPNVKTITDREGLALLDIEFENYADGKVPYIKMALLESILRRF 899
Db 829 GNVHRLAARNVLSRPSQVADFCVADLLPRDDQQLYSKAPPIKMALESIHFGY 888
Qy 900 THQSDVWSYGVYVWELMTFGAKPYDGI PAREIPDLLEKGRILPOPICITIDVYMIWVKCM 959
Db 889 THQSDVWSYGVYVWELMTFGAERYAGRLAEVDPDLLEKGRILAQPOICITIDVYMWVKCM 948
Qy 960 MIDSECRPRFRLLVSEFSHMADPORFVVIQNEIDLGA---SPLDSTFYRSLLEDODMGD 1016
Db 949 MIDENIRPTFKELANFTRMADPRRYLVIKRES--GPGIAPGEBPHGLTKKLEVELEBP 1007
Qy 1017 LVDABEYLVPOGFCPPDDAPGA---GGMVHHRHSSSTRS-----CGSL 1059
Db 1008 ELDDLDLAEEDNLATTTGLSALIPVGTINRPGSOSILSPSSGYMPOAGNLGSGSQ 1067
Qy 1060 TLGLEPSEBEADPSPLA---PSEGASDVFDGDLGMAKGIQS-----LPTHDS 1107
Db 1068 ESAVGSSEBRCP-RPVSLHMRGCLASSESSGCHV-TGSEAELOKRVMSCRSRSRSPR 1125
Qy 1108 P-----LQRY-S-BDPTVPL-----PSETDGYVAPLTCSPPEYVNOPRVPO-PPSP 1152
Db 1126 PGDSAYHQRSHSLTPVTPPLSPGLEBEDVNGYV-----MPDTHLKTPTSS 1172
Qy 1153 RCGPLPAARPAATLERPKTLSPGKGVV-----KDVAFGCAVENPEYILPPOGGAAPQ 1206
Db 1173 REGTIS-----SVGLSSVLTGTEEBEDD-----EEYNNRRRRHSP- 1208
Qy 1207 PHPAPFAPAFDNLVYWD-----ODPEKGAPEPTFKGTALENPEYL 1249
Db 1209 PHPAPRSSLEELGYEMDVGSLSLGSSTGSCPLHPVIMPTAGTTPDEDEYEM 1263

```

```

; TITLE OF INVENTION: RECEPTOR AND ERBB RECEPTORS
; FILE REFERENCE: 49321-136
; CURRENT APPLICATION NUMBER: US/11/113,202
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893
; PRIOR FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-113-202-23

Query Match 27.5%; Score 1874.5; DB 7; Length 419;
Best Local Similarity 89.6%; Pred. No. 2.9e-112;
Matches 353; Conservative 5; Mismatches 29; Indels 7; Gaps 2;

Qy 1 MELAALCRWGLLALLPREGASTOYCTGDMKRLPASPETHLMDLRHLVGGCOVVOGML 60
Db 1 MELAALCRWGLLALLPREGASTOYCTGDMKRLPASPETHLMDLRHLVGGCOVVOGML 60
Qy 61 ELTYLPTWASLSFLQDIEVQGYVLIAHQVQVPLQRLIVRGTLPEEDNYALAVLDNG 120
Db 61 ELTYLPTWASLSFLQDIEVQGYVLIAHQVQVPLQRLIVRGTLPEEDNYALAVLDNG 120
Qy 121 DPLNNTPTVTSAGPGLRELOLRSLTEILKGGVLIQNNPOLCYODTILMKDIFHNQOLA 180
Db 121 DPLNNTPTVTSAGPGLRELOLRSLTEILKGGVLIQNNPOLCYODTILMKDIFHNQOLA 180
Qy 181 LTLIDTNSRACHPSPCKSGRCWGESSESDQSILTRYVCAGGACRCGRLPTDCHBQC 240
Db 181 LTLIDTNSRACHPSPCKSGRCWGESSESDQSILTRYVCAGGACRCGRLPTDCHBQC 240
Qy 241 AAGCTGPKSDCLACLFHNSGICELHCPALVTYNTDIFESMNPBEGRTTFGASCVTACP 300
Db 241 AAGCTGPKSDCLACLFHNSGICELHCPALVTYNTDIFESMNPBEGRTTFGASCVTACP 300
Qy 301 YNYISTDVSGCTLVCPRLHNOEVTAEADGTORCEKSKPCARVCYGLAMEHLREVAVTSAN 360
Db 301 YNYISTDVSGCTLVCPRLHNOEVTAEADGTORCEKSKPCARVCYGLAMEHLREVAVTSAN 360
Qy 361 IOEPAGCKKIFGSLAFLBESPDGASNTAPLOP 394
Db 356 LMQPG--PAHPVLSFLRPSMDVLSAFYSLPLAP 387

RESULT 12
US-11-113-202-4
; Sequence 4, Application US/11/113,202
; Publication No. US20050272637A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; APPLICANT: Shantleh, Lara
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1
; FILE REFERENCE: 49321-136
; CURRENT APPLICATION NUMBER: US/11/113,202
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 419
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: Herstatin; receptor non-binding mutant (Arg to Ile mutation at

```

```

; OTHER INFORMATION: residue 371)
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (371)..(371)
; OTHER INFORMATION: position
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; US-11-113-202-4
Query Match      27.5%; Score 1873.5; DB 7; Length 419;
Best Local Similarity 89.6%; Pred. No. 3.4e-112;
Matches 353; Conservative 5; Mismatches 29; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLRPASPEYHLDMLRHLVGGCCVVGQNTL 60
DB 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLRPASPEYHLDMLRHLVGGCCVVGQNTL 60
QY 61 ELYTPTNASLSFLDIOEVGGVYLIAHQVQVPLQRLIVRGTLFEDNYALAVLNG 120
DB 61 ELYTPTNASLSFLDIOEVGGVYLIAHQVQVPLQRLIVRGTLFEDNYALAVLNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSTLTKGGLVLIQRNPOLCYODTILMKDIFHKNQOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSTLTKGGLVLIQRNPOLCYODTILMKDIFHKNQOLA 180
QY 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGCGARCKGRLPTDCCHQC 240
DB 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGCGARCKGRLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTYNTDFPSMNPDEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTYNTDFPSMNPDEGRYTFGASCVTACP 300
QY 301 YNYLSTDVGSCTLVCPPLHNQVETABDGTQRCCKSKPCARVYGLGMEHLREVRVTSAN 360
DB 301 YNYLSTDVGSCTLVCPPLHNQVETABDGTQRCCKSKPCARVYGLGMEHLREVRVTSAN 360
QY 361 IOEPAGCKKIFGSLAFPLPSPFGDASNTAPLQIP 394
DB 361 IOEPAGCKKIFGSLAFPLPSPFGDASNTAPLQIP 394
QY 394 LRMQPG--PAHPVLSFLIPSWDLVSAFYSPLAP 387
DB 394 LRMQPG--PAHPVLSFLIPSWDLVSAFYSPLAP 387

RESULT 13
US-11-113-202-2
; Sequence 2, Application US/11113202
; Publication No. US20050272637A1
; GENERAL INFORMATION:
; APPLICANT: Clinton, Gail M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING SIGNALING BY IGF-1
; TITLE OF INVENTION: RECEPTOR AND ERBB RECEPTORS
; FILE REFERENCE: 49321-136
; CURRENT APPLICATION NUMBER: US/11/113,202
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/590,473
; PRIOR FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US 60/564,893
; PRIOR FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (342)..(342)
; OTHER INFORMATION: Applicants herein disclose Thr and Ser sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (345)..(345)
; OTHER INFORMATION: Applicants herein disclose Leu and Pro sequence variants at this
; OTHER INFORMATION: position
; FEATURE:

```

```

; NAME/KEY: MISC FEATURE
; LOCATION: (346)..(346)
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (356)..(356)
; OTHER INFORMATION: Applicants herein disclose Leu and Gln sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (357)..(357)
; OTHER INFORMATION: Applicants herein disclose Arg and Cys sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (358)..(358)
; OTHER INFORMATION: Applicants herein disclose Met and Leu sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (361)..(361)
; OTHER INFORMATION: Applicants herein disclose Gly, Asp, Ala and Val sequence
; OTHER INFORMATION: variants at this position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (371)..(371)
; OTHER INFORMATION: Applicants herein disclose Arg and Ile sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (376)..(376)
; OTHER INFORMATION: Applicants herein disclose Leu and Ile sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (394)..(394)
; OTHER INFORMATION: Applicants herein disclose Pro and Arg sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (404)..(404)
; OTHER INFORMATION: Applicants herein disclose Pro and Leu sequence variants at this
; OTHER INFORMATION: position
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (413)..(413)
; OTHER INFORMATION: Applicants herein disclose Asp and Asn sequence variants at this
; OTHER INFORMATION: position
; US-11-113-202-2
Query Match      27.3%; Score 1860.5; DB 7; Length 419;
Best Local Similarity 88.8%; Pred. No. 2.3e-111;
Matches 350; Conservative 3; Mismatches 34; Indels 7; Gaps 2;

QY 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLRPASPEYHLDMLRHLVGGCCVVGQNTL 60
DB 1 MELAALCRWGLLALLPPGAASSTQVCTGDMKRLRPASPEYHLDMLRHLVGGCCVVGQNTL 60
QY 61 ELYTPTNASLSFLDIOEVGGVYLIAHQVQVPLQRLIVRGTLFEDNYALAVLNG 120
DB 61 ELYTPTNASLSFLDIOEVGGVYLIAHQVQVPLQRLIVRGTLFEDNYALAVLNG 120
QY 121 DPLNNTTPTVGTASPGGLRELQRLSTLTKGGLVLIQRNPOLCYODTILMKDIFHKNQOLA 180
DB 121 DPLNNTTPTVGTASPGGLRELQRLSTLTKGGLVLIQRNPOLCYODTILMKDIFHKNQOLA 180
QY 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGCGARCKGRLPTDCCHQC 240
DB 181 LTLIDTNSRACHPCSPMKSGRCWGESSEDCQSLTRTVCAAGCGARCKGRLPTDCCHQC 240
QY 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTYNTDFPSMNPDEGRYTFGASCVTACP 300
DB 241 AAGCTGPKHSDCLAHFNHSGICELHCPALVTYNTDFPSMNPDEGRYTFGASCVTACP 300

```

Db 241 AAGCTPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACP 300
Qy 301 VNYLSTDVSSCTLVNCEPLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSAN 360
Db 301 VNYLSTDVSSCTLVNCEPLHNOEYTAEDGTORCEKSKPCAR---GXHXXPRDAVVPVXX 357
Qy 361 IOEPAGCKKIFGSLAFLPESFDGDPASNTAPLP 394
Db 358 XQXPRA---HPVLSFLKSMXVSAFYSLPLAP 387

RESULT 14
US-11-077-386-25

Sequence 25, Application US/11077386
Publication No. US20050272067A1
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Turner, Leah R.
APPLICANT: Chen, Huel-Mei
APPLICANT: Rodriguez, Maria
APPLICANT: Liu, Shu-Hui
TITLE OF INVENTION: Compositions, Splice Variants and Methods Relating to Cancer Spec
FILE REFERENCE: DEX-0537
CURRENT APPLICATION NUMBER: US/11/077,386
CURRENT FILING DATE: 2005-03-10
PRIOR APPLICATION NUMBER: US 60/566,706
PRIOR FILING DATE: 2004-04-30
PRIOR APPLICATION NUMBER: US 60/565,144
PRIOR FILING DATE: 2004-04-23
PRIOR APPLICATION NUMBER: US 60/551,911
PRIOR FILING DATE: 2004-03-10
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.1
SEQ ID NO 25
LENGTH: 534
TYPE: PRT
ORGANISM: Homo sapien
US-11-077-386-25

Query Match 15.5%; Score 1056.5; DB 7; Length 534;
Best Local Similarity 41.1%; Pred. No. 3.4e-60;
Matches 223; Conservative 76; Mismatches 202; Indels 41; Gaps 13;
Qy 10 GLLALLPPGAA--STVCTGTGDMKRLRPASPEYTHDMLRHYGCGOVVQGNLELYPT 67
Db 11 GLLFELARGSEVGNQAVCPPTLNGLSVYGAENQYOTLYKLYERCEVVMGMLLEIVTGH 70
Qy 68 NASLSFLDIOEVOGVVLIHNOVQVPLQRLRIYRGQLFEDNYALAVLDNGDPLNNTT 127
Db 71 NADLSFLQWIREVTVLVANNEFSTLPLNLRVVRGTVQVQDKKFAIFVM-----LNTYT 125
Qy 128 PVTGASPGGLRELQRLSLEILKGGVLIQRPOLCYOTIIMKDIFFHKNQALATLIDTN 187
Db 126 ----NSSHALQRLRLQTLLEILSGVYIEKNKLCMNTIDMRDIVRBD---AELVVD 178
Qy 188 RSRACHPSPMCKGRCKWGESSEDCQSLTRIVCAGGC-ARCKGRLPTDCCHCQCAAGCTG 246
Db 179 NGRSCPCHGEVCKG-RCWGPGESEDCQTLTKITICAPQCNHCGFPNPNOCCHDECGAGGSG 237
Qy 247 PHSQCLACLFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACPYNVYST 306
Db 228 PDDTCFACRHHNDSGACVPRCPQLVYNKLTFLQLEPNPHTKYQYGVGVASCPHNFV-V 296
Qy 307 DVGSCITLVCPHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSANIQEPAG 366
Db 297 DQTSQVRACPDPKMEVD-KNGLKMCEPCGGLCPKACBEGTSG-SRFQTVSSNIDGFVN 353
Qy 367 CKKIFGSLAFLPESFDGDPASNTAPLPQPOLQVFEILEITIGVLYISAMPDLPULSVQ 426
Db 354 CTKIIGNLDPLITGLNGDPMHKI PALDPEKLVFRLVREITIGLYNLQSPMPHMHFSVS 413
Qy 427 NLQVIRGRILHNGAVS-LTLQGLGISWLGRLSRLBELSGLALIHNTHLCPHTVPWDL 485

Db 414 NLTTIGSLVYRGFSLIMKLVNTSLGPFSLKELISAGRIYISANROLCYHSHSINTKV 473
Qy 466 FRNPHQALLHRA-NRPEBCVGE-----GLACHOLCARHCMPPT 526
Db 474 LRGPTEERLDIKHNRPRDC-GECKGLLGENRESGRRLKGLFCPRGRSVEGNNQGG 532
Qy 527 QC 528
Db 533 GC 534

RESULT 15
US-11-077-386-26

Sequence 26, Application US/11077386
Publication No. US20050272067A1
GENERAL INFORMATION:
APPLICANT: Macina, Roberto
APPLICANT: Turner, Leah R.
APPLICANT: Chen, Huel-Mei
APPLICANT: Rodriguez, Maria
APPLICANT: Liu, Shu-Hui
TITLE OF INVENTION: Compositions, Splice Variants and Methods Relating to Cancer Spec
FILE REFERENCE: DEX-0537
CURRENT APPLICATION NUMBER: US/11/077,386
CURRENT FILING DATE: 2005-03-10
PRIOR APPLICATION NUMBER: US 60/566,706
PRIOR FILING DATE: 2004-04-30
PRIOR APPLICATION NUMBER: US 60/565,144
PRIOR FILING DATE: 2004-04-23
PRIOR APPLICATION NUMBER: US 60/551,911
PRIOR FILING DATE: 2004-03-10
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.1
SEQ ID NO 26
LENGTH: 400
TYPE: PRT
ORGANISM: Homo sapien
US-11-077-386-26

Query Match 11.7%; Score 798; DB 7; Length 400;
Best Local Similarity 43.2%; Pred. No. 6.1e-44;
Matches 163; Conservative 53; Mismatches 141; Indels 20; Gaps 9;
Qy 10 GLLALLPPGAA--STVCTGTGDMKRLRPASPEYTHDMLRHYGCGOVVQGNLELYPT 67
Db 11 GLLFELARGSEVGNQAVCPPTLNGLSVYGAENQYOTLYKLYERCEVVMGMLLEIVTGH 70
Qy 68 NASLSFLDIOEVOGVVLIHNOVQVPLQRLRIYRGQLFEDNYALAVLDNGDPLNNTT 127
Db 71 NADLSFLQWIREVTVLVANNEFSTLPLNLRVVRGTVQVQDKKFAIFVM-----LNTYT 125
Qy 128 PVTGASPGGLRELQRLSLEILKGGVLIQRPOLCYOTIIMKDIFFHKNQALATLIDTN 187
Db 126 ----NSSHALQRLRLQTLLEILSGVYIEKNKLCMNTIDMRDIVRBD---AELVVD 178
Qy 188 RSRACHPSPMCKGRCKWGESSEDCQSLTRIVCAGGC-ARCKGRLPTDCCHCQCAAGCTG 246
Db 179 NGRSCPCHGEVCKG-RCWGPGESEDCQTLTKITICAPQCNHCGFPNPNOCCHDECGAGGSG 237
Qy 247 PHSQCLACLFHNSGICELHCPALVTYNTDFESMPNDEGRYTFGASCVTACPYNVYST 306
Db 228 PDDTCFACRHHNDSGACVPRCPQLVYNKLTFLQLEPNPHTKYQYGVGVASCPHNFV-V 296
Qy 307 DVGSCITLVCPHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSANIQEPAG 366
Db 297 DQTSQVRACPDPKMEVD-KNGLKMCEPCGGLCPKACBEGTSG-SRFQTVSSNIDGFVN 353
Qy 367 CKKIFGSLAFLPESFDG 383
Db 354 CTKIIGNLDPLITGLNG 370

Wed Jan 4 10:12:16 2006

us-09-930-125-2.rapbn

Page 12

Search completed: January 3, 2006, 11:28:36
Job time : 20 secs
